A Novel Multimodal Approach for Studying the Dynamics of Curiosity in Small Group Learning

Tanmay Sinha, Zhen Bai, Justine Cassell

1

A Novel Multimodal Approach for Studying the Dynamics of Curiosity in Small Group Learning

Tanmay Sinha[1], Zhen Bai[2], Justine Cassell[3]

Abstract—Curiosity is a vital metacognitive skill in educational contexts, leading to creativity, and a love of learning. And while many school systems increasingly undercut curiosity by teaching to the test, teachers are increasingly interested in how to evoke curiosity in their students to prepare them for a world in which lifelong learning and reskilling will be more and more important. One aspect of curiosity that has received little attention, however, is the role of peers in eliciting curiosity. We present what we believe to be the first theoretical framework that articulates an integrated socio-cognitive account of curiosity that ties observable behaviors in peers to underlying curiosity states. We make a bipartite distinction between individual and interpersonal functions that contribute to curiosity, and multimodal behaviors that fulfill these functions. We validate the proposed framework by leveraging a longitudinal latent variable modeling approach. Findings confirm a positive predictive relationship between the latent variables of individual and interpersonal functions and curiosity, with the interpersonal functions exercising a comparatively stronger influence. Prominent behavioral realizations of these functions are also discovered in a data-driven manner. We instantiate the proposed theoretical framework in a set of strategies and tactics that can be incorporated into learning technologies to indicate, evoke, and scaffold curiosity. This work is a step towards designing learning technologies that can recognize and evoke moment-by-moment curiosity during learning in social contexts and towards a more complete multimodal learning analytics. The underlying rationale is applicable more generally for developing computer support for other metacognitive and socio-emotional skills.

Index Terms—Curiosity, Learning in Social Contexts, Multimodal Human Behavior Analyses, Scaffolding

I. INTRODUCTION

CURIOSITY is defined as a strong desire to learn or know more about something or someone, and is an important metacognitive skill to prepare students for lifelong learning [1]. Because curiosity is increasingly rare in students in classrooms, perhaps because increasingly rare in curricula [2], [3], large-scale attempts exist to help teachers bring back children's curiosity[1]. Traditional accounts of curiosity in psychology and neuroscience focus on how it can be evoked via underlying mechanisms such as novelty (features of a stimulus that have not yet been encountered), surprise (violation of expectations), conceptual conflict (existence of multiple incompatible pieces of information), uncertainty (the state of being uncertain), and

anticipation of new knowledge [4], [5]. These knowledge-seeking experiences create a positive impact on student's beliefs about their competence in mastering scientific processes, in turn promoting greater breadth and depth of exploration [6].

Such theories have inspired the development of computer systems for facilitating task performance by enhancing an individual's curiosity (e.g., [6], [7], [8]), and simulating human-like curiosity in autonomous agents [9]. Evoking curiosity here mainly focuses on directing an individual to a specific new knowledge component, followed by facilitating knowledge acquisition through exploration. Such a linear approach largely ignores how learning is influenced by social contexts. Here, a child's intrinsic motivation, exploratory behaviors, and subsequent learning outcomes may be informed not only by materials available to the child, but also the active work of other children, and the presence of facilitators [10], [11]. For instance, an expression of uncertainty or a hypothesis made by one child may cause peers to realize that they too are uncertain about the topic under discussion. This may initiate working together to overcome the cause of uncertainty, in turn having a positive impact on the individuals' curiosity [12].

Although substantial literature exists on the intrapersonal origins of curiosity, with rare exceptions (cf. [3], [13]) a lacuna exists concerning how peer-driven social factors contribute to moment by moment changes in curiosity. Peer influence is important because even in classrooms where teachers teach to the test, group work is encouraged, and such participation structures remain a core part of classrooms [10]. Thus, peers may be one of the few ways that curiosity is elicited. This makes it critical to understand curiosity beyond the individual level to an integrated knowledge-seeking phenomenon shaped by the individual, physical and social environment. Embodied Conversational Agents (ECAs) - particularly child-like ECAs called *virtual peers* - have demonstrated a special capacity in supporting learning and collaborative skills for young children [14]. Knowing how social factors influence curiosity allows the design of ECAs and other technologies to support curiosity-driven learning before children naturally support each other.

To address the goal of a framework that accounts for both individual and social influences on curiosity, our work takes a novel approach. We look at the impact of individuals on other individuals in group work by focusing on the observable behaviors that peers use - both language and nonverbal behavior - and using latent variable modeling to connect those observable behaviors to the students' inner curiosity states. Our work is unique in moving from a theoretical model to building arguments and evidence, relying on machine learning and human annotation, concerning verbal and

[1]T. Sinha is with Professorship for Learning Sciences and Higher Education, ETH Zürich, Switzerland. E-mail: tanmay.sinha@gess.ethz.ch

[2]Z. Bai is with Department of Computer Science, University of Rochester, USA. E-mail: zbai@cs.rochester.edu

[3]J. Cassell is with School of Computer Science, Carnegie Mellon University, USA and Paris Artificial Intelligence Research Institute, Inria, France. E-mail: justine@cs.cmu.edu

[1]https://tinyurl.com/TeachTeachersToEvokeCuriosity

nonverbal indicators that curiosity exists in a given student (and verbal and nonverbal behaviors that are most successful in eliciting curiosity, to building a statistical model that tests the theoretical assumptions.

More specifically, we propose an integrated socio-cognitive account of curiosity based on literature spanning psychology, learning sciences and group dynamics, and empirical analyses of a group informal learning environment. We make a bipartite distinction between functions that contribute to curiosity, and multimodal behaviors that fulfill these functions. Such functions are latent and therefore not directly knowable, and we must therefore hypothesize or presume them based on observable phenomena that index them. We therefore refer to them as putative functions. These functions comprise (i)"knowledge identification and acquisition (helps humans realize that there is something they desire to know, and leads to the acquisition of the desired new knowledge), and (ii) "knowledge intensification" (escalates the process of knowledge identification or acquisition by providing favorable environment) - at the individual and interpersonal level. We perform statistical validation of this framework to illuminate predictive relationships between multimodal behaviors, functions (latent variables because they cannot be directly observed), and ground truth curiosity (as judged by naive annotators). Longitudinal latent variable modeling is used to explicitly account for group structure and differentiate fine-grained behavioral dynamics.

The main contributions of this work are three-fold: First, it begins to fill the research gap of how social factors, especially interpersonal peer dynamics in group work, influence curiosity (section 6.B, 6.C). Second, the model is designed to lay a theoretical foundation to inform the design of learning technologies, a virtual peer in the current study, that can employ pedagogical strategies to evoke and maintain curiosity in social environments (section 6.C). Findings derived from the current analyses of human-human interaction can be informative in guiding the learning technology (human-agent interaction) design (section 7.A), and in decision making for multimodal analyses of behavior (section 7.B). Third, at the methodological level, our research (i) introduces novel approaches for collecting rich multimodal data in group settings (section 3.A), which is key to making fine-grained behavioral inferences, (ii) advances the use of crowdsourcing platforms for efficient ground truth annotation, which is important in human behavior analysis for educational research and beyond (section 5.A), (iii) provides a rigorous and reproducible semi-automatic behavior annotation approach, which combines complementary strengths of state-of-the-art machine learning methods and advantages of human judgment (section 5.B, 5.C and 5.D).

In what follows, Section 2 discusses related work. Section 3 discusses data collection (pre-studies and main study) across multiple learning contexts. Section 4 describes our combination of theory-driven and data-driven process for the development of the theoretical framework of curiosity. Section 5 describes the annotation of ground truth curiosity, verbal and nonverbal behaviors, and turn-taking metrics. Section 6 discusses statistical validation of the proposed theoretical framework of curiosity, with a discussion of the model fit to our corpus, identified causal interpersonal and intrapersonal behavior influence patterns that result in increase or maintenance of curiosity, and initial implementations of our theory-and-data driven approach in a virtual peer that is a co-player of a science game with children. In section 7, we describe the implications of this work for learning technology design and multimodal learning analytics. We end with limitations, future work and conclusion in sections 8, 9 and 10.

II. RELATED WORK

For clarity, we divide related work into discussions about curiosity in the psychology, group dynamics and learning sciences literature and end with a brief overview of existing computational modeling approaches.

A. Curiosity in the Psychology Literature

Researchers in psychology describe curiosity as a psychological and behavioral state that "responds to an inconsistency or a gap in knowledge" [15], and raises "feelings of mystery, of strangeness, and of wonder" [16]. Like hunger and thirst, curiosity is considered a critical internal drive for human beings that causes us to explore our environment, acquire knowledge and learn skills. It is generally described as an intrinsically motivated desire, passion or appetite for information, knowledge and learning [11], [17]. Several theoretical lenses explain the cause of curiosity. For instance, the incongruity theory argues that curiosity arises from the human tendency to make sense of the world on observing violated expectations [18], [19]. Theories of conflict arousal consider curiosity as a drive that leads to the simultaneous occurrence of incompatible response tendencies. Main determinant factors for psychological conflict include both perceptual factors (e.g., novelty, surprise etc) and epistemic factors (e.g., incongruity, confusion etc) [20].

The information-gap perspective, another dominant view in the field, proposes that curiosity is raised when people attend to a gap in their knowledge, and the intensity depends on importance, salience and surprisingness of the desired information [21], [22], [4]. Other research in psychology has conceived of curiosity as desirable feelings associated with the anticipation of acquiring new knowledge. As described further below, our perspective adds to these causal factors verbal challenges, incompatible hypotheses, and negative evaluations of the target child's ideas Previous literature also reveals several behavioral cues for curiosity. These include physical exploration such as orientation, locomotion and manipulation towards objects of interest, epistemic investigation such as question asking, experimentation, reasoning about observed phenomena, as well as expressions of surprise, excitement, wonder, confusion and attentiveness [20], [23], [24].

Existing behavior-based measures of children's curiosity mainly rely on physical exploration and simple verbal behaviors such as question asking and commenting (e.g., [25], [26]). There is, however, a lack of rigorous operational measures of curiosity that incorporate the verbal and nonverbal behaviors displayed in real-time interaction.

B. Curiosity in the Group Dynamics Literature

Social accounts for curiosity remain largely unexplored in the psychology literature, and relevant research primarily

focuses on parent-child or teacher-student interaction [27], [3], instead of peer-peer interaction. Knowledge dissonance, social comparison and social information-seeking are three closely-knit factors related to curiosity and exploration which have been studied in prior work. First, compared to teacher-led learning where a teacher holds higher power/status in terms of knowledge possession, peer-peer learning is more likely to result in challenging different opinions or ideas from one another and active resolution of such knowledge incongruity or dissonance [28]. This is one of the main sources of curiosity [29]. Controversy, instantiated as conflict or disagreement, has also been identified as one of the social causes of dissonance - the simultaneous existence of cognitions that in one way or another do not fit together. When individuals working in a small group experience dissonance, they might work towards reducing or eliminating dissonance via different means. One important means is attempt to seek additional social support for the held opinion by emphasizing its importance, lucrativeness, etc., in turn triggering curiosity [28].

Second, through social comparison, students are more eager to evaluate the correctness of their own opinion via group discussion and coming to know their peers' opinions, compared to knowing an expert's opinion [30]. When individuals face a question with no clear solution and they cannot reduce the uncertainty by consulting objective sources of information, they turn to views endorsed by others in the group and evaluate the accuracy of their beliefs by comparing themselves to others [28]. Also, cognitive and affective changes are more likely when observing others who are perceived as friends or similar to the observer [10]. Also, students are more likely to actively seek information and solutions when their uncertainty is shared or at least considered as warranted, reasonable, or legitimate by their peers [12]. Such joint hardship [31] is likely to impact group member's behavior positively, due to the trigger it provides for engaging in cooperative/joint effort to overcome the obstacle by reasoning or physical exploration.

Third, social information seeking, or a general interest in gaining new social information (how others behave, act and feel) promotes acceptance (a non-evaluative feeling with unconditional positive regard towards another) [31] and creates mutually shared cognition in the group [32]. It creates space for group members to learn from others' preferences and viewpoints [33]. Increased group member familiarity and knowledge awareness can increase willingness to work jointly and lead to consideration of more alternatives. Our research is meant to fill the research gap that exists concerning the role of individuals in social contexts, by incorporating theories from peer learning and group dynamics. Note that this perspective is quite separate from and complementary to more holistic accounts of group work, such as Cress' [34] account of the co-evolution of collaborative knowledge, that describes the interplay between individual learning and group learning and its impact on group-level curiosity. Curiosity at the level of the group is surely at work in the contexts we analyzed as well, however, our focus on verbal and nonverbal behaviors produced by individuals and affecting other individuals' underlying psychological states precluded an analysis at the level of the group.

C. Curiosity in the Learning Sciences Literature

Discussions about curiosity can be traced back to [35], who differentiated extensive curiosity (that widens a learner's interest) and particular curiosity (that helps a learner acquire detailed knowledge). [36] tied these notions to the literature on knowledge awareness in collaborative learning settings. They posited that when a learner's activities are oriented towards the "same knowledge" that the peer is looking at, discussing or changing, particular curiosity is excited. We find, on the contrary, that when a learner's activities are oriented towards "different knowledge" than their peer, extensive curiosity can be evoked and collaboration possibilities are enhanced.

Curiosity has also been discussed under the umbrella of intrinsic motivation [37]. Intrinsically motivated learners derive pleasure from the task itself, while learners with extrinsic motivation rely on external rewards. [38] considers curiosity as a motivational aspect in the design of learning technologies, and discusses surprising students as a central instructional tactic to lead them to explore new areas of the subject for constructing coherent explanations. Curiosity is also proposed as one of seven dimensions of the construct of "learning power" [39], which refers to a form of consciousness, or critical subjectivity leading to growth. Critically curious learners are comparatively less likely to accept what they are told uncritically, and more willing to reveal their questions and uncertainties. It is important to note, however that, the focus of the learning sciences literature has been fundamentally cognitive, whereas we seek to understand the social scaffolding of curiosity along with its cognitive roots.

In the field of artificial intelligence in education, there is scarce research on how the dynamics of social interaction may influence student's intrinsic learning motivation. Some applications of curiosity can be seen in the development of pedagogical agents. For example, [40] found that a curious peer (that keeps questioning) can problematize the interaction, and direct learner's attention to spot the contradiction in their knowledge structure, thereby inducing curiosity. [41] also discovered that if a computer agent displays curiosity by pro-actively responding to novel conflicting stimuli, it can discover interesting learning concepts. Interaction with such a computer agent was shown to lead human learners to engage in greater task exploration of the learning environment, enhanced attention, and improved learning outcomes. The underlying model of interpersonal influence that is common to both these research strands is "modeling". These systems implicitly assume human learners to spontaneously pick up on social cues of the "curious" pedagogical agent. We aim to develop a more nuanced understanding of curiosity, without directly equating behaviors to curiosity, or, relying solely on a single theoretical lens of looking at curiosity.

D. Computational Models of Curiosity

In general, curiosity has been computationally modeled [42] using an appraisal process where the incoming stimuli are first evaluated for their potential to provide an appropriate stimulation level. Subsequently, the stimulation level is mapped to a non-linear emotion curve called the Wundt curve [43] for deriving the curiosity level. The Wundt curve postulates too

little stimulation to result in boredom, too much stimulation to result in anxiety, and only an optimal stimulation level to result in curiosity. We, however focus not just on perceptual, but also on the epistemic dimension of curiosity. Further, our computational model is meant to serve as an input into a human-pedagogical agent interaction in group work, unlike the goals of prior computational models of curiosity.

III. DATA CORPUS

To capture the simultaneous dynamic of fine-grained behaviors of multiple participants (Figure 1), a novel human behavior recording apparatus was developed. Recording equipment comprised four Logitech C920 Pro Webcam devices for the egocentric recording (frontal view), eight camcorder recorders for the exocentric recording (single and group view), and four lapel-microphones (Sennheiser ew100 G2 and Samson UHF Micro 32) to separate individual's speaking (Figure 2). We were able to get a frontal view of participants for automated detection of facial expression, head and gaze gesture, and single view and group view for analyzing other individual and social behaviors (Figure 3). Our goal in collecting such rich multimodal data was to be able to then annotate ground truth curiosity, detect verbal and nonverbal behaviors that fulfill putative functions of curiosity (see section 4.A, 4.B), and subsequently, statistically verify predictive relationships between them (see section 6).

A. Pre-study 1 (In-School)

Twenty participants in 6th grade (usually 11 to 12 years old) participated in our first pre-study. Participants were recruited by teachers in a local public school. There was no remuneration. The study took place in the 6th grade science class and STEAM class for two weeks. There were eight observation sessions in total, four sessions for science class and four sessions for STEAM class, with each session lasting about 45 minutes. The science class was focused on the topic of earth science learning, while the main activities of the STEAM class included robotics, programming, and crafting. Both classes were divided into small groups with 3-4 students. There were several constraints for data collection in a classroom, including large-scale observation (3-4 groups of students at the same time), noisy environments, and confined space and time for equipment set up. To accommodate the above constraints, we used a lightweight recording setup that included two camcorder recorders to capture students from the opposite sides of the table, and one lapel microphone in the middle of the table. Researchers also took field notes to describe student's interpersonal learning behaviors relating to curiosity.

B. Pre-study 2 (Science Summer Camp)

Thirty-one participants aged 9-14 participated in our second pre-study. Participants were recruited by teachers. The study took place in two science summer camp sessions hosted in a local child maker-space, with each summer camp lasting one week. We observed 30 hours in total. There was a wide variation of hands-on activities in the summer camp, spanning topics from physics, chemistry, biology, life science, robotics, and crafting. The class was divided into four to five groups.

Compared to pre-study 1, the summer camp observation was longer and noisier, with more diverse (and less structured) activities. The class layout also changed frequently. We used one camcorder recorder and one lapel-microphone fixed on the table for each group of children. Similar to pre-study 1, researchers took field notes to describe peer-peer interaction associated with curiosity.

C. Main Study (In-Lab)

Our main study, which built upon the two pre-studies, was conducted in a semi-controlled adult-free lab setting. It comprised forty-four participants in 5th or 6th grade (16 male and 28 female, aged between 10-12, average age 11.2), who collaboratively built a Rube Goldberg machine (RGM) for about 35-40 minutes. The RGM task included building several chain reactions meant to be triggered automatically to trap a ball in the cage (without external human support). The machine was created using a variety of simple objects such as rubber bands, pipe cleaners, toy cars, clothespin, etc. We employed the RGM building activity because it supports learning of key science knowledge for students in 5th and 6th grade as defined by the Pennsylvania Department of Education [44] such as force, motion and energy transfer, and enables collaborative hands-on learning and creative problem solving [45] without need for adult intervention.

Participants were recruited through flyers sent to local public schools, parent mailing-lists of the university, and advertisements on social media and in public spaces. All participants were remunerated with $30 cash. 12 study sessions were videotaped, each one including a group of three to four students, and lasting for one and half hours. The procedure for each study session was as follows: (1) 10 minutes for an ice-breaking game; (2) 5 minutes free exploration of materials easily found at home; (3) 5 minutes of introducing participants to Rube Goldberg machines (RGM); (4) 30 minutes for collaboratively building the RGM; (5) a 5-minute opportunity to demonstrate the built RGM; (6) a 10-minute interview, and optionally (7) 5 minutes free exploration of a pre-made RGM; (8) 20 minutes for dispensing remuneration.

IV. THEORETICAL FRAMEWORK DEVELOPMENT

We take a socio-cognitive view of curiosity whereby we acknowledge social influences but also try to isolate the individual mind as a cognitive unit of analysis by controlling for these external influences. This theoretical stance is partly borne out of our interest in studying individual curiosity (as opposed to a construct that is defined at the dyadic or group level like rapport). Adopting such a view also serves the practical goal of using behavioral cues of the other individuals to influence the curiosity of a given individual.

We define curiosity-driven learning in social contexts to be situations where particular forms of interaction among people trigger or facilitate behaviors associated with high curiosity. To leverage this notion for developing an integrated psychological and social framework of curiosity, a combination of theory-driven and data-driven approaches was used. From the theory-driven perspective, several iterations of the literature review were conducted with a gradual shift from individual- to

Fig. 1. Empirical observation across learning contexts. Left: in-lab RGM building; Middle: in-school STEAM class; Right: science summer camp

Fig. 2. In-lab study data collection apparatus. Left: equipment arrangement; Middle: real arrangement; Right: fixture of four Webcam devices.
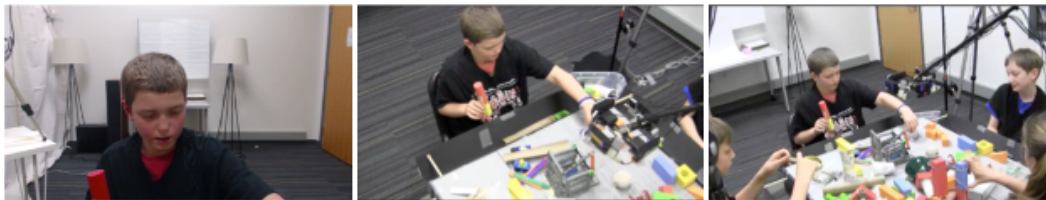

Fig. 3. Illustration of video recording in the In-lab study. Left: frontal view; Middle: single view; Right: group view.

interpersonal-level curiosity. Starting with research on curiosity at the individual level in psychology, we fanned out to critically investigate existing research on the social influence of curiosity within the group dynamics and the Learning Sciences literature. Informed by these theoretical lenses, we developed (i) a set of putative functions that contribute to curiosity, and (ii) multimodal behaviors that provide evidence for the potential presence of an individual's curiosity in the current time-interval because they fulfil these functions.

From the data-driven perspective, we carried out empirical observation of small groups of 9-14 year old children engaged in hands-on learning activities across lab, science and STEAM class, and informal learning environments (section 3). Qualitative analyses were conducted in an exploratory manner first based on notes, videos and audio data collected during field observations [46]. The pre-studies helped in identifying a list of curiosity-related individual and interpersonal behaviors during small group science learning, based on our initial literature review. Building on these diverse pre-study contexts, we conducted follow-up qualitative/quantitative analyses on our in-lab data (but not constrained by the initial behavior list), to empirically validate and extend curiosity-related behaviors while minimizing confirmation bias.

Overall, this mixed-methods approach involving the classroom, science summer camp, and lab as study contexts allowed us to (i) obtain familiarity to describe and explain the phenomenon of curiosity in different small group learning contexts, (ii) develop the integrated theoretical framework

of curiosity with in-depth empirical evidence, and (iii) form relevant hypotheses based on the empirical exploration for follow-up quantitative validation of the theoretical framework.

We originally hoped to use both the pre-studies and main study data for quantitative analyses, but then discovered that adults in the classroom and summer program systematically obviated contexts in which curiosity might appear. The adults asked leading questions, rhetorical questions, and other kinds of classroom moves that left little room for exploration, creativity and curiosity. That in itself is an interesting finding, but one which we would like to address in subsequent work. While it is often said that classrooms represent more ecologically valid learning situations than do labs, we were interested in peer-peer interaction without adult intervention. Thus, in the current work, we looked at contexts in which curiosity might naturally occur in the absence of adults and examined the interpersonal triggers of that curiosity. For that reason, we also designed a lab context with open-ended activities where children were free to behave as they wished in the absence of adult supervision. This, we believe, was as close to ecologically valid as possible, given the presence of video cameras and adult (non-participating) observers.

A. Putative Functions Contributing to Curiosity

The iterative process described above led to the emergence of three function groups at the individual and interpersonal level. Because curiosity has traditionally been described as an inherently individual and stable disposition toward seeking novelty and approaching unfamiliar stimuli, we first outline

TABLE I
CORPUS EXAMPLES OF BEHAVIOR SEQUENCES. P1 IS THE CHILD WITH HIGH CURIOSITY (SEE SECTION 5.A FOR HOW WE ANNOTATED CURIOSITY)

Behavior Cluster	Empirical Observation (Example 1)	Empirical Observation (Example 2)
Cluster 1,2	**P1**: Hey let's..wait I have an idea *[idea verbalization]* **P1**: Let's see what this is, but let me just, let me just.. *[proposes joint action, co-occurs with physical demonstration, initiates joint inquiry]* **P2**: I have no idea how to do this, but it's making my brain think *[positive attitude towards task]*	**P1**: So the chain has to be like this *[idea verbalization with iconic gesture]* **P1**: How would that be? *[question asking followed by orienting towards stimulus]* **P1**: Well, I don't want it to break, so I want it to be about...no, let's say half an...half an inch *[causal reasoning to justify actions being taken]*
Cluster 1,3	**P1**: Wait we need to raise it a bit higher *[making suggestions]* **P1**: Maybe if we put it on..Umm..this thing maybe..this is high enough? *[co-occurs with joint stimulus manipulation]* **P2**: Why? W-Why do we need to make it that high? *[disagreement and asking for evidence]*	**P2**: And the funnel can drop it into one of um..those things **P1**: If the funnel can drop it **P1**: Okay but then..even if it hits this, then we need what is this going to hit? *[challenge]* **P1**: Here- let- just- make sure that it's going to hit it *[followed by physical demonstration/verification]*
Cluster 2,3,4	**P1**: Roll off into here and go in there *[hypothesis generation]* **P1**: Okay, so how are we going to do that? *[question asking]* **P2**: It looks like something should hit the ball *[making suggestion]*	**P2**: We could use this if we wanted *[making suggestion]* **P1**: Let's figure this quickly...so we at least have this part done *[preceded by expression of surprise and followed by trying to connect multiple objects to create a more complex object]*

individual aspects of curiosity for each function. We then flesh out interpersonal aspects of curiosity for every function. Each function can be realized in several different behavioral forms.

We call the first function group *Knowledge Identification*. As curiosity arises from a strong desire to obtain new knowledge that is missing or doesn't match with one's current beliefs, a critical precondition of this desire is to realize the existence of such knowledge. At an *individual* level, knowledge identification contributes to curiosity by increasing awareness of gaps in knowledge [21], and highlighting relationships with related or existing knowledge to assimilate new information [47]. Further, exposure to novel and complex stimulus can raise uncertainty, subsequently resulting in conceptual conflict [20], [29]. At an *interpersonal* level, knowledge identification contributes to curiosity by developing awareness of somebody else in the group having conflicting beliefs [20] and awareness of the knowledge they possess [36], so that a shared conception of the problem can be developed [32].

We call the second function group *Knowledge Acquisition*. Knowledge-seeking behaviors driven by curiosity not only contribute to the satisfaction of the initial desire for knowledge but also potentially lead to further identification of new knowledge. For example, question asking may help close one's knowledge gap by acquiring desired information from another peer. Depending on the response received, however, it may also lead to escalated uncertainty or conceptual conflict relating to the original question, thus reinforcing curiosity. At an *individual* level, knowledge acquisition involves finding sensible explanation and new inference for facts that do not agree with existing mental schemata [48], [47], and can be indexed by the generation of diverse problem-solving approaches [48], [49]. It also comprises comparison with existing knowledge or search for relevant knowledge through external resources to reduce simultaneous opposing beliefs stemming from the investigation [28]. At an *interpersonal* level, knowledge acquisition comprises revelation of uncertainties in front of group members [39], joint creation of new interpretations and ideas, engagement in argument to reduce dissonance among peers [50], and critical acceptance of what is told [39].

Finally, we call the third function group *Intensification of Knowledge Identification and Acquisition*. The intensity of curiosity or the desire for new knowledge is influenced by factors such as the confidence required to acquire it [21], its

incompatibility with existing knowledge and existence of a favorable environment [28]. At an *individual* level, intensification of knowledge identification and acquisition can stem from factors such as anticipation of knowledge discovery [31], interest in the topic [38], willingness to try out tasks beyond ability without fear of failure [49], taking ownership of own learning and being inclined to see knowledge as a product of human inquiry [39]. These factors can subsequently result in a state of increased pleasurable arousal [20]. At an *interpersonal* level, intensification of knowledge identification and acquisition is influenced by the willingness to get involved in group discussion and the tendency to be part of a cohesive unit in pursuit of instrumental objectives and/or for the satisfaction of a group member's affective needs [28].

Such willingness can span from the spectrum of merely continuing interacting to pro-actively reacting to the information others present [32]. Various interpersonal factors play out along this spectrum. Salient ones include interest in knowing more about a group member [51], promotion of unconditional non-evaluative regard towards them [31], the tendency of spontaneous pickup of behavior initiated by a group member (where the initiator did not display any communicated intent of getting the others to imitate) [28], and awareness of one's own uncertainty being shared or considered legitimate by those peers [12]. All these factors can result in a cooperative effort to overcome common blocking points for the group [31].

B. Behaviors Fulfilling Putative Functions of Curiosity

Behavioral episodes including language and associated multimodal communicative signals (e.g., facial expressions, gaze) serve as both communicative markers, i.e they provide evidence for the presence of curiosity of group members, and mind markers, i.e they shape group member's understanding and expectation of how to approach the task, along with conceptualization and construction of the associated knowledge [52]. They can (i) contain a single action or multiple co-occurring or contingent actions made by one or more individuals, (ii) be purposeful or non-purposeful because the underlying human strategy that governs the sequence of behaviors is unknown. Our review of prior research in psychology and learning sciences led us to link the behaviors with their putative functions in evoking curiosity, and organize these behaviors into four clusters. Table 1 illustrates examples of these behavior clusters from our empirical observations.

Cluster 1 corresponds to behaviors that enable an individual to get exposed to and investigate physical situations, which may spur socio-cognitive processes beneficial to curiosity-driven learning [20], [47]. Examples include orientation (using eye gaze, head, torso) and interacting with stimuli (e.g., manipulation of objects). When looking at video segments tagged with high curiosity in our empirical data, these behaviors occur in contexts where children look at different aspects of the stimulus (e.g., the function of novel objects, physical properties of mineral samples in the science class, transition phase of dry ice samples in the summer camp, etc) by orienting towards it using their gaze and torso, smelling or scratching it, rotating and trying to fit more than one object together, etc. *Cluster 2* corresponds to behaviors that enable an individual to actively make meaning out of their observations [20], [53], [47]. Examples include idea verbalization, justification and generating hypotheses. *Cluster 3* corresponds to behaviors that involve joint investigation with other peers [20], [53], [47]. Examples include arguing, evaluating the problem-solving approach of a partner (positive or negative), expressing disagreement, making suggestions, sharing findings, question asking. In video segments tagged with high curiosity, these behaviors occur in contexts where children listen to other's suggestions, express disagreement or challenge their perspective by pointing out loopholes, and engage in a physical demonstration for clarification. Finally, *Cluster 4* corresponds to behaviors that reveal an individual's affective states [54], [11] including surprise, enjoyment, confusion, uncertainty, flow and sentiment towards task.

We hypothesize that behaviors across these clusters will map onto one or more putative functions of curiosity, because there can be many different functions or reasons why a communicative behavior occurs. For example, in knowledge-based conflict in group work, attending to differing responses of others compared to one's own may raise simultaneous opposing beliefs *(knowledge identification)*. This awareness might in turn activate cognitive processes, wherein an individual may seek social support for one's original belief by emphasizing its importance and validating ones idea by providing justification, or, engaging in a process of back and forth reasoning to come to a common viewpoint *(knowledge acquisition)*. Further, this awareness may as well impact socio-emotional processes, where an individual may perceive a conflict differently and their emotions felt and expressed might vary depending on relation with and perception of the source of conflict, e.g., is it a more competent/less competent, more cooperative/less cooperative peer that raises conflict, and therefore take the next action of resolving that conflict differently *(intensification of knowledge identification and acquisition)*. We intend to discover prominent mappings between functions described in section 4.A and behaviors described in section 4.B more formally in a data-driven way in section 6.

V. QUANTITATIVE ANALYSES OF IN-LAB STUDY

We now describe fine-grained quantitative analyses from a convenience sample of the first 30 minutes (out of 35-40 minutes given each group), of the RGM task (lab study) for half of the sample; that is, 22 children across 6 groups. Table 2 provides a summary of all coding metrics used in this article. Our goal is to empirically verify the theoretical framework of curiosity proposed in section 4.

A. Assessment of Ground Truth Curiosity

Person perception research has demonstrated that judgments of others based on brief exposure to their behaviors give an accurate assessment of interpersonal dynamics [55]. We used Amazon's Mechanical Turk (MTurk) platform to obtain ground truth for curiosity via such a thin-slice approach, using the definition "curiosity is a strong desire to learn or know more about something or someone", and a rating scale comprising 0 (not curious), 1 (curious) and 2 (extremely curious). Crowdsourcing platforms provide benefits of a diverse sample of raters who can be accessed quickly, easily and for relatively little cost [56]. Our previous research has successfully deployed thin-slice coding for other social phenomena such as interpersonal rapport in peer tutoring using MTurk [57], [58].

Here, four naïve raters annotated every 10-second slice of videos of the interaction for each child presented in randomized order. We post-processed ratings by removing raters who used less than 1.5 standard deviation time compared to the mean time taken for all rating units (HITs). We then computed a single measure of Intraclass correlation coefficient (ICC) for each possible subset of raters for a particular HIT, and then picked ratings from the rater subset that had the best reliability for further processing. Finally, inverse-based bias correction [59] was used to account for label overuse and underuse, and to pick one single rating of curiosity for each 10-second thin-slice. The average ICC of 0.46 aligns with reliability of curiosity in prior work [60], [61].

B. Assessment of Verbal Behaviors

We adopted a mix of semi-automatic and manual annotation procedures to code 11 verbal behaviors, in line with the curiosity-related behavioral set described in section 4.B. These verbal behaviors span propositional [62] and interpersonal [63] functions of contributions to a conversation. Propositional functions are those that are fulfilled by contributing informational content to the dialog (e.g., idea verbalization, justification), and interpersonal functions are those that are fulfilled by managing the relationship between the interlocutors (e.g., social question asking, positive evaluation). Five verbal behaviors were coded using a semi-automatic approach - *uncertainty, argument, justification, suggestion* at the clause level, and *agreement* at the turn level. A clause contains a subject (a noun or pronoun) and a predicate (conjugated verb that says something about what the subject is or does). During a full turn, a speaker holds the floor and expresses one or more interpretable clauses.

First, a particular variant of neural language models called paragraph vector or doc2vec [64] was used to learn distributed representations for a clause/turn. This means that for every clause/turn in our data corpus, we transformed the sequence of words in it to a tuple (or vector) of continuous-valued features that characterize the semantic meaning of those words. Such feature representation implies that sentences in a test set that functionally similar to sentences in a training set can still

TABLE II

Summary of coding methods (detailed coding scheme at http://www.tinyurl.com/codingschemecuriosity)

Construct	Definition used to code/infer the construct	Coding method
Ground Truth Curiosity	A strong desire to learn or know more about something or someone.	Four MTurk raters annotated each 10-sec thin slice; average ICC = 0.46; used inverse-based bias correction to pick the final rating.
Verbal Behavior		
1. Uncertainty	Lack of certainty about ones choices or beliefs, and is verbally expressed by language that creates an impression that something important has been said, but what is communicated is vague, misleading, evasive or ambiguous. e.g - *"well maybe we should use rubberbands on the foam pieces"*, *"wait do we need this thing to funnel it through?"*	Used a semi-automated annotation approach: after automatic labeling of these verbal behaviors, two trained raters (Krippendorff's alpha > 0.6) independently corrected machine annotated labels; average percentage of machine annotation that remained the same after human correction was 85.9 (SD = 12.71).
2. Argument	A coherent series of reasons, statements, or facts intended to support or establish a point of view. e.g -*"no we got to first find out the chain reactions that it can do"*, *"wait, but anything that goes through is gonna be stuck at the bottom"*	
3. Justification	The action of showing something to be right or reasonable by making it clear. e.g -*"oh we need more weight to like push it down"*, *"wait with the momentum of going downhill it will go straight into the trap"*	
4. Suggestion	An idea or plan put forward for consideration. e.g - *"you could kick a ball to kick something"*, *"you are adding more weight there which would make it fall down"*	
5. Agreement	Harmony or accordance in opinion or feeling; a position or result of agreeing. e.g - *"But we need to have like power, and weight too"* (Quote) — *"Yeah we need more weight on this side"* (Response), *"And we put the ball in here..I hope it still works, and it goes..so it starts like that, and then we hit it"* (Quote) — *"Ok that works"* (Response)	
6. Question Asking (On-Task/Social)	Asking any kind of questions related to the task or non-task relevant aspects of the social interaction. e.g - *"so what's gonna..what will happen like after the balls gets into the cup?"*, *"why do we need to make it that high?"*, *"do you want to build something like a chain reaction or something like that?"*, *"do you two go to the same school?"*, *"who else watched the finale of gravity falls?"*	Used manual annotation procedure due to unavailability of existing training corpus (Krippendorff's alpha > 0.76 between two raters).
7. Idea Verbalization	Explicitly saying out an idea, which can be just triggered by an individual's own actions or something that builds off of other peer's actions. e.g - *"yeah that ball isn't heavy enough"*, *"so it's like tilted a bit up so it catches it instead of tilted down"*	
8. Sharing Findings	An explicit verbalization of communicating results, findings and discoveries to group members during any stage of a scientific inquiry process. e.g - *"look how I'm gonna see I'm gonna trap it"*, *"look I made my pillar perfect"*	
9. Hypothesis Generation	Expressing one or more different possibilities or theories to explain a phenomenon by giving relation between two or more variables. e.g - *"we could use scissors to cut off the baby's head which would cause enough friction"*, *"okay we need to make it straight so that the force of hitting it makes it big"*	
10. Task Sentiment (Positive/Negative)	A view of or attitude (emotional valence) toward a situation or event; an overall opinion towards a subject matter. We were interested in looking at positive or negative attitude towards the task that students were working on. e.g - *"oh it's the coolest cage I've ever seen, I'd want to be trapped in this cage"*, *"ok so I'm gonna try to find out a way for the end to make this one go and fall"*,*"I'm getting very mad at this cage"*,*"but I don't know how to make it better"*	
11. Evaluation (Positive/Negative)	Characterization of how a person assesses a previous speaker's action and problem-solving approach. It can be positive or negative. e.g - *"oh that's a pretty good idea - that was a good idea"*,*"let's make this thing elevated and make it go down"*,*"oh wait this doesn't- you're not pushing anything over here"*, *"no it can't go like that otherwise it will be stuck"*	
Nonverbal Behavior (AU - facial action unit)		
1. Joy-related	AU 6 (raised lower eyelid) *and* AU 12 (lip corner puller).	Used an open-source software OpenFace for automatic facial landmark detection, and a rule-based approach post-hoc to infer affective states
2. Delight-related	AU 7 (lid tightener) *and* AU 12 (lip corner puller) *and* AU 25 (lips part) *and* AU 26 (jaw drop) *and not* AU 45 (blink).	
3. Surprise-related	AU 1 (inner brow raise) *and* AU 2 (outer brow raise) *and* AU 5b (upper lid raise) *and* AU 26 (jaw drop).	
4. Confusion-related	AU 4 (brow lower) *and* AU 7 (lid tightener) *and not* AU 12 (lip corner puller).	
5. Flow-related	AU 23 (lip tightener) *and* AU 5 (upper lid raise) *and* AU 7 (lid tightener) *and not* AU 15 (lip corner depressor) *and not* AU 45 (blink) *and not* AU 2 (outer brow raise).	
6. Head Nod	Variance of head pitch.	Used OpenFace to extract head orientation, and computed variance post-hoc
7. Head Turn	Variance of head yaw.	
8. Lateral Head Inclination	Variance of head roll.	
Turn Taking		
1. Indegree	A weighted product of number of group members whose turn was responded to (*activity*) and total time that other people spent on their turn before handing over the floor (*silence*).	Used two novel metrics constructed using an application of social network analysis for weighted data.
2. Outdegree	A weighted product of number of group members to whom floor was given to (*participation equality*), and the amount of time spent when holding floor before allowing a response (*talkativeness*).	

achieve good predictions. The motivation for this approach stems from: (i) lack of available corpora of verbal behaviors that are large enough, and collected in similar settings as ours (groups of children engaged in scientific inquiry), and hence (ii) limited applicability of traditional n-gram based models to cross-domain settings, which would result in a high-dimensional representation with poor semantic generalization, (iii) limitations of other neural language models such as word2vec that do not explicitly represent word order and surrounding context (in contrast, doc2vec models contain an additional paragraph token that acts as a memory and remembers what is missing from the current context, thus not ignoring word order that is important for the semantic representation), and (iv) our desire to reduce manual annotation time when each child's behaviors must be annotated.

Based on the recommended procedure in [64], we used concatenated representations of two fixed size vectors of size 100 that we learned for each sentence as input to a machine learning classifier (L2 regularized logistic regression) - one learned by the standard paragraph vector with distributed memory model, and one learned by the paragraph vector with a distributed bag of words model. Empirically too, we found this concatenated vector representation to perform better on cross-validation performance on the training data, compared to using any of the two vector representations alone. Training data for the five verbal behaviors annotated using this process is shown in the right column of Table 3, along with standard performance metrics such as weighted F1 score (to account for class imbalance) and Area under ROC curve (AUC). Test data comprised the in-lab study corpus (section 4.C).

The robustness of machine annotated labels on the test data was ensured using human annotators. Two raters first coded presence or absence of verbal behaviors on a random sample of 100 clauses/turns following a coding manual given to them for training, and computed inter-rater reliability using Krippendorff's alpha. Once raters reached a reliability of > 0.7 after one or more rounds of resolving disagreements, they independently rated a different set of 50 clauses/turns independently, and we computed the final reliability on these (left column of Table 3, and > 0.6 for all behaviors). Subsequently, the raters independently de-noised or corrected machine annotated labels for the full corpus, and we use these final labels for empirical validation of the theoretical framework (as described in section 6). Compared with this human ground truth, the average ratio of false positives to false negatives across all annotation categories in the machine prediction was 14.18 (SD = 12.31), suggesting over-identification of the presence of verbal behaviors.

TABLE III

RESULTS FROM SEMI-AUTOMATIC VERBAL BEHAVIOR ANNOTATION. RIGHT COLUMN DESCRIBES EXTERNAL CORPUS USED FOR TRAINING MACHINE LEARNING CLASSIFIERS & THEIR PERFORMANCE. LEFT COLUMN DEPICTS INTER-RATER RELIABILITY FOR HUMAN JUDGMENT USED IN VERIFYING ROBUSTNESS OF MACHINE ANNOTATED LABELS.

Verbal Behavior [Krippendorff's α for human judgment]	Training Data for Semi-Automated Classification {Weighted F1, AUC} (10-fold cross validation)]
1. Uncertainty [0.78]	Wikipedia corpus manually annotated for 3122 uncertain 7629 certain instances [65] {0.695, 0.717}
2. Argument [0.792]	Internet Argument Corpus manually annotated for 3079 argument and 2228 non argument instances [66]. Argument quality score split at 70% to binarize class label {0.658, 0.706}
3. Justification [process (0.936), causal (0.905), model (0.821), example (0.731), definition (0.78), property (0.847)]	AI2 Elementary Science Questions corpus manually annotated for 6 kinds of justification - process, causal, model, example, definition, property [67]. Reported performance is the average performance of 6 binary machine learning classifiers {0.766, 0.696}
4. Suggestion [0.608]	Product reviews [68] and Twitter [69] corpuses manually annotated for 1000 explicit suggestion and 13000 explicit non-suggestion instances {0.938, 0.865}
5. Agreement [0.935]	LiveJournal forum and Wikipedia discussion corpuses manually annotated for 2754 agreement and 8905 disagreement instances based on quote and response pairs [70] {0.717, 0.696}

We found that the most common false positives were cases where a clause or turn comprised one word (e.g., okay), backchannels (e.g., - hmmm..) and very short phrases lacking enough context to make a correct prediction. The average percentage of machine annotated labels that did not change even after the human de-noising step was 85.9 ($SD = 12.71$). This meant that majority of the labels were correctly predicted in the first place. This was also reflected in a good cross validation training performance of the models (right column of Table 3). Six other verbal behaviors (*question asking (on-task, social) ($\alpha = 1$), idea verbalization ($\alpha = 0.761$), sharing findings ($\alpha = 1$), hypothesis generation ($\alpha = 0.79$), attitude towards task (positive, negative) ($\alpha = 0.835$), evaluation sentiment (positive, negative) ($\alpha = 0.784$)*) were coded using manually due to unavailability of existing training corpus.

C. Assessment of Nonverbal Behaviors

Our current work focused initially on the nonverbal behaviors of the face, operationalized as facial landmarks, and based on a rich body of literature describing movements of the face in learning [71], and particularly the facial action units that index certain emotions that often co-occur with curiosity [60]. This work has discovered consistent associations (correlations as well as predictions) between particular facial configurations and human emotional or mental states [54], [72], [60]. We used automated visual analysis to construct five feature groups corresponding to emotional expressions that provide evidence for the presence of the affective states of *joy, delight, surprise, confusion* and *flow* (a state of engagement with a task such that concentration is intense). A simple rule-based approach was followed (see Table 2) to combine emotion-related facial landmarks, which were previously extracted on a frame by frame basis using a state-of-the-art open-source software OpenFace [73]. We then selected the most dominant (frequently occurring) emotional expression for every 10-second slice of the interaction for each group member, among all the frames in that time interval.

Automated visual analysis was also used to capture variability in head angles for each child in the group, which correspond to *head nods (i.e. pitch), head turns (i.e. yaw)*, and *lateral head inclinations (i.e. roll)*. The motivation for using head movement in our curiosity framework is inspired by our prior work in nonverbal behavior and learning technologies [74] and in multimodal analytics [75], [76] which have emphasized the contribution of nonverbal cues in inferring behavioral constructs such as interest and involvement that are closely related to the construct of curiosity, and which have demon-

strated the positive impact of nonverbal behavior by virtual peers in children's learning. By using OpenFace[73], we first performed frame by frame extraction of head orientation and then calculated the variance post-hoc to capture intensity in head motions for every 10 second of the interaction for each group member. Because head pose estimation takes as input facial landmark detection, we only considered those frames that had a face tracked and facial landmarks detected with confidence greater than 80%.

D. Assessment of Turn Taking Dynamics

While the annotated verbal behaviors fulfill propositional and interpersonal conversational goals in the social interaction, the interactional function of contributions to a conversation is captured by turn-taking behaviors. Interactional discourse functions are "responsible for creating and maintaining an open channel of communication between the participants" [62]. The motivation for capturing turn taking in the current research stems Further from prior literature that has used measures such as participation equality and turn taking freedom as indicators of involvement in small-group interaction [77].

In the current work, we designed two novel metrics using a simple application of social network analysis for weighted data. By representing speakers as nodes and time between adjacent speaker turns as edges, we computed two features for each group member. These features calculated for every 10 seconds, comprised (i) *TurnTakingIndegree*, a weighted product of the number of group members whose turn was responded to (activity) and the total time that other people spent on their turn before handing over the floor (silence), and was quantified as $activity^{1-\alpha} * silence^{\alpha}$. Because high involvement is likely to be indexed by higher activity and lower silence, α was set to -0.5, (ii) *TurnTakingOutdegree*, a weighted product of the number of group members to whom floor was given (participation equality), and the amount of time spent when holding floor before allowing a response (talkativeness), and was quantified by $participation\ equality^{1-\alpha} * talkativeness^{\alpha}$. Because higher participation equality and talkativeness are favorable, α was set to +0.5. These two metrics were used in our empirical validation.

VI. EMPIRICAL VALIDATION OF THE PROPOSED THEORETICAL FRAMEWORK OF CURIOSITY

We used a "multiple-group" version of continuous time structural equation model (CTSEM) [78] to evaluate the proposed theoretical framework of curiosity, and statistically verify the predictive relationships between ground truth curiosity (formalized as a manifest variable), putative functions

described in our theoretical framework (formalized as latent variables) and multimodal behaviors (formalized as time-dependent predictors). For clarity, we first present a rationale for the choice of our modeling framework, and then a formal description of the technical background underlying these latent variable analyses. Finally, we describe a concrete application of this modeling framework to our in-lab study corpus.

A. Modeling Framework (CTSEM)

1) Rationale: Conventional Structural Equation Models (SEMs) assume independence of observations, and thus cannot be applied directly to analyze autocorrelated time series data from multimodal behavior analyses. This suggests consideration of a Dynamic Bayesian Network-like model to explicitly model temporal dependencies between the latent random variables across time-steps. Applications of such models in the social and behavioral sciences are usually limited to discrete time models, with the assumption that time progresses in discrete steps, and intervals between measurement occasions are equal. In many cases, these assumptions are not met, resulting in biased parameter estimates and a misunderstanding of the strength and time course of effects. Continuous time SEM models overcome these limitations by using multivariate stochastic differential equations to estimate an underlying continuous process and recover underlying latent or hidden causes linking the entire sequence.

2) Technical Description: To clarify how CTSEM improves over SEM models in the current context, we give a short technical description here. Formally, a multivariate stochastic differential equation for a latent process of interest in CTSEM can be written as $d\eta_i(t) = A\eta_i(t) + Bz_i + M\chi_i(t) + GdW_i(t) + \xi_i$ *(Structural part of the SEM model)*, where A is the drift matrix that models auto effects the latent variable has on itself on the diagonals, and cross effects to other latent processes on the off-diagonals, in turn characterizing temporal relationships between the processes. ξ_i determines the long-term level of the latent process. Matrix B and M represent the effect of time-independent and time-dependent variables on the latent process. Time-independent predictors would typically be variables that differ between subjects, but are constant within subjects for the time range in question (e.g., a trait curiosity questionnaire). Time-dependent predictors vary over time and are independent of fluctuations of the latent processes in the system. They can be treated as a simple impulse form where the predictors are treated as impacting these processes only at the instant of an observation.

The matrix G represents the effect of noise or the stochastic error term $dW_i(t)$ on the change in the latent process. $Q = G * G^T$ represents the variance-covariance matrix of diffusion process in continuous time. The essence of diffusion processes is to capture very slow patterns of change in the latent variable. Further, this latent process can be used to predict manifest variables of interest using the equation $y_i(t) = \Lambda\eta_i(t) + \zeta_i(t)$ *(Measurement part of the SEM model)*, where Λ is a matrix of factor loadings between the latent and manifest variables and ζ_i is the residual (error) vector.

A Kalman filter can be used to fit CTSEM to the data and obtain standardized estimates for the influence of behaviors on latent functions, and in turn these latent functions on curiosity. It uses a series of measurements observed over time (containing statistical noise and other inaccuracies) to produce estimates of unknown variables that tend to be more precise than those based on a single measurement alone. It is a state space model described by a (i) state equation that describes how the latent states change over time and is analogous to *structural part of the SEM model*, and (ii) output equation that describes how the latent states relate to the observed states at a single point in time (how the observed output is produced by the latent states), and is analogous to *measurement part of the SEM model*. In the presence of multiple groups in a dataset (e.g., we have 6 groups in our corpus), a "multiple group" version of CTSEM should be used. It allows investigation of group level differences and helps understand variability in model parameters across different groups.

3) Application to the In-Lab Study Corpus: Because knowledge identification and acquisition are closely intertwined with knowledge-seeking behaviors and it is hard to distinguish between these underlying mechanisms based on observable or inferred multimodal behaviors, we formalized them under the same latent variable. The final set of latent functions that we statistically verified therefore included: (i) *individual* knowledge identification and acquisition, (ii) *interpersonal* knowledge identification and acquisition, (iii) *individual* intensification of knowledge identification and acquisition, (iv) *interpersonal* intensification of knowledge identification and acquisition. We ran two versions of CTSEM.

In the first version, we specified a model where only factor loadings between the manifest variable and latent variables in *measurement part of the model* were estimated for each group distinctly (we report the average and standard deviation across the 6 groups in Figure 4), but all other model parameters including those belonging to *structural part of the model* were constrained to equality across all groups (Model$_{constrained}$) and then estimated freely. This means that matrices A, B, M, G and Λ were freely estimated. Because the form of a behavior does not uniquely determine its function, nor vice-versa, we did not pre-specify the exact pattern of relationships between behaviors and functions to look for/estimate. In the second version of the model, all parameters for all groups were estimated distinctly (Model$_{free}$).

Technically, the rich representational capacity of "multiple-group" CTSEM allows running these two separate models. However, analytically, the decision to separately run these two models was based on the intuition that while the relationships between the appearance of behaviors and their contribution to the latent functions of curiosity would remain the same across groups, the relative contribution of interpersonal or individual tendencies for knowledge identification, acquisition and intensification would vary based on learning dispositions of people towards seeking the unknown. This intuition stemmed from prior literature that has looked into measuring learning dispositions [39], an important dimension of which is the ability of learners to balance between being sociable and being private in their learning, i.e not being completely independent or dependent, but rather working interdependently. We hypothesized that this dimension will impact curiosity
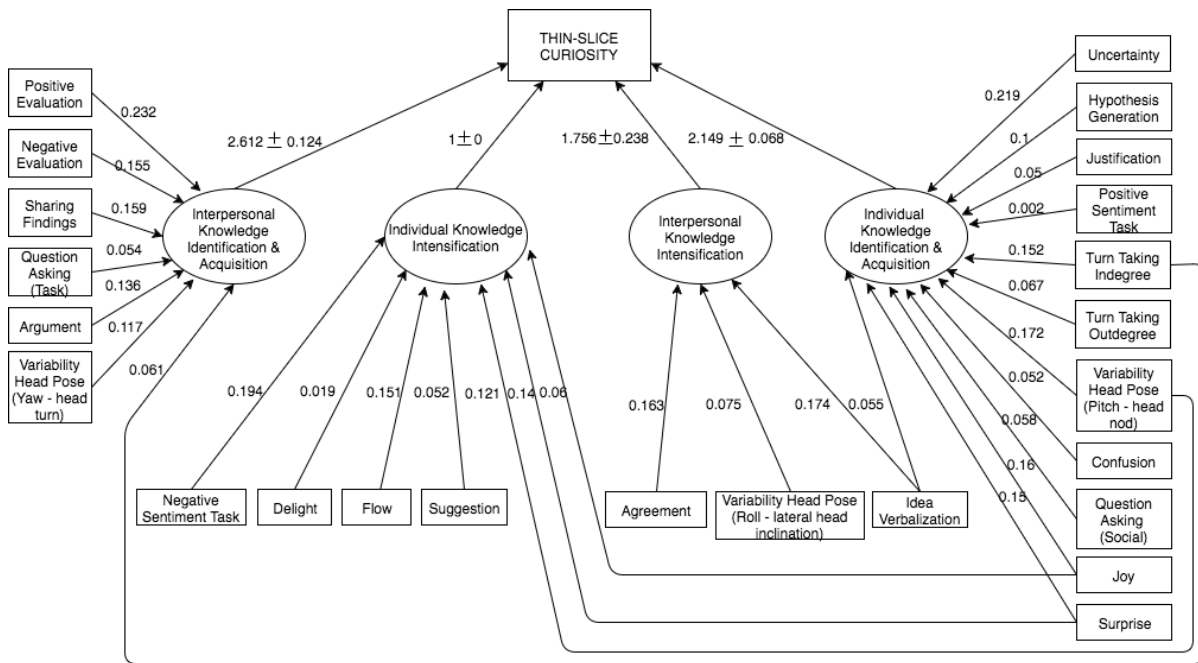
Fig. 4. Empirical validation of the theoretical framework of curiosity depicting fit of the Continuous time structural equation model (CTSEM) to the In-Lab study corpus. Rectangles represent observed constructs, while ovals represent latent constructs. Direction and degree of predictive influences are represented by edges between multimodal behaviors (time dependent predictors), putative functions of curiosity (latent variables) and thin-slice curiosity (manifest variable). Degree of predictive influence between latent and manifest variable is averaged across 6 study groups.

differently, especially because our data comprised a group learning context, and therefore expected $Model_{constrained}$ to fit the data better than $Model_{free}$.

B. Model Results and Discussion

An empirical validation confirmed our hypothesis of $Model_{constrained}$ fitting to the data better than $Model_{free}$. The Akaike Information Criterion (AIC) for $Model_{constrained}$ (933.48) was ∼3x lower than $Model_{free}$ (2278.689). We now illustrate results of the CTSEM ($Model_{constrained}$) in Figure 4, depicting links with top-ranked standardized estimates between behaviors and latent variables. In few cases, we also added links with the second-highest standardized estimate if they clarified our interpretation of the latent function. Because a latent variable is unmeasured, its units of measurement must be fixed by a researcher for the model to be identified (for there to exist a unique solution for all of the model's parameters). Therefore, one factor loading was conventionally fixed to 1.

Overall, these results confirm the correctness of our framework along three main aspects: (i) The grouping of behaviors under each latent function and their contribution to individual and interpersonal aspects of knowledge identification, acquisition and intensification aligns with prior literature on the intrapersonal origins of curiosity, but also teases apart the underlying interpersonal mechanisms, (ii) There exists strong and positive predictive relationships between these latent variables and thin-slice curiosity, (iii) Knowledge identification and acquisition have a stronger influence to curiosity than knowledge intensification, and interpersonal-level functions have stronger influence compared to individual-level functions. We now discuss latent functions and associated behaviors, ordered by the degree of positive influence on curiosity.

First, *Interpersonal Knowledge Identification and Acquisition* shows the strongest influence on curiosity among the four latent functions (2.612 ± 0.124). The natural merging of knowledge identification and knowledge acquisition corroborates with the notation that one person's knowledge-seeking may draw the attention of another group member to a related knowledge gap and escalate collaborative knowledge-seeking. Behaviors that positively contribute to this function are mainly from cluster 3 (*sharing findings, task-related question asking, argument, and evaluation of other's ideas*). Also, nonverbal behaviors including *head turn* and *turn taking dynamics (indegree)* are also related to this function, which support the idea that a higher degree of group members' interest and involvement in the social interaction stimulates awareness of peer's ideas, subsequently leading to knowledge-seeking via social means to gain knowledge from the experience of others, and add that onto one's own direct experiences.

Second, *Individual Knowledge Identification and Acquisition* shows a strong influence on curiosity (2.149 ± 0.066). Similar to the interpersonal level function, knowledge identification and acquisition merge into one coherent function, as knowledge-seeking behaviors can sparkle new unknown or conflicting information within the same individual. Behaviors from cluster 2 (*hypothesis generation, justification, idea verbalization*) and cluster 4 (*confusion, joy, surprise, uncertain, positive sentiment towards task*) mainly contribute to this function. *Head nod*, indicative of positive feelings towards the stimulus due to its compatibility with the response [79], maps to this function as well. Finally, *turn taking (indegree and outdegree)* and *social question asking* contribute positively to individual knowledge identification and acquisition. Interest in others reflects a general level of trait curiosity [51].

Third, we find that a relatively small group of behaviors including *agreement, idea verbalization* and *lateral head inclination* have a predictive influence on the latent function of *Interpersonal Knowledge Intensification*, which in turn has a high positive influence on curiosity (1.756 ± 0.238). Agreement may contribute to information seeking by promoting acceptance and cohesion. Working in social contexts broadcasts idea verbalization done by an individual to other group members, which might in turn increase their willingness to get involved. Lateral head inclination during the RGM activity is associated with an intensive investigation of the RGM solution offered by both oneself and other group members. Overall, engagement in a cooperative effort to overcome common blocking points in the group work may result in intensifying knowledge-seeking.

Finally, the latent function of *Individual Knowledge Intensification* has the least comparative influence on curiosity. It is associated with nonverbal behaviors such as *head nod* and emotional expressions of positive affect (*flow, joy* and *delight*), which function towards increasing pleasurable arousal. Also, *surprise* and *suggestion* positively influence this latent function and signal increased anticipation to discover novelty, conceptual conflict, and correctness of one's idea. Interestingly, results also show that *negative sentiment about the task* positively influences an individual's knowledge-seeking behaviors. Qualitative corpus observations reveal that such verbal expressions often co-occur with the evaluation made by a group member within the same 10-second thin-slice, signaling a desire for cooperation. Thus, a potential explanation of this association is that expressing negative sentiment about the task may signal hardship, which draws group member's attention and increases chances of receiving assistance, thus increasing engagement in knowledge-seeking.

C. Causal Intra/Inter-personal Influence Patterns

The CTSEM analyses described above establish the overall degree of predictive influence among verbal/nonverbal behaviors, latent functions and ground truth thin-slice curiosity. However, they do not provide "causal" insights into inter-and-intra-personal curiosity dynamics of the kind that can be leveraged by a learning technology to foster curiosity in group work. We therefore conducted follow-up causal analyses with the dual goal of generating such actionable insights and providing additional empirical evidence for our proposed theoretical framework of curiosity.

Specifically, we ran conditional Granger causality [80] to assess the interdependence among verbal/nonverbal behaviors and their causal contribution to the increase ($0 \rightarrow 1$, $0 \rightarrow 2$, $1 \rightarrow 2$) or maintenance ($1 \rightarrow 1$, $2 \rightarrow 2$) of the thin-slice curiosity level (or vice-versa). This notion of causal influence is based on the idea that if the variance of the autoregressive prediction error of time-series A at the present time is reduced by inclusion of past measurements from time-series B, then B is said to have a causal influence on A. Because such a causal relation (based on cause-effect relations with constant conjunctions) between A and B can be direct, mediated by a third time-series C, or be a combination of both, the technique of conditional Granger causality allows modeling causal relationship among

multivariate behavioral time series. More technical details can be found in [81].

Here we report four categories of causal influence patterns in Table 4. First, the top-left corner highlights the causal influence of other's past behavior (within the last minute) on one's increase/maintenance of curiosity. Such triggering behaviors span both verbal (e.g., justification, idea verbalization, task question asking, negative evaluation, etc) and nonverbal (e.g., head movement variability, confusion and surprise-related facial expressions) categories. Second, the top-right corner highlights the reciprocal pattern of causal influence, i.e., other's increase/maintenance of curiosity in the past (within the last minute) on one's behaviors. Such triggered behaviors include not just uncritical acceptance in the form of a head nod or verbal agreement, but also those that signal investment into furthering knowledge acquisition like on-task question tasking, justification and flow-related facial expressions. Taken together, these bi-directional sets of interpersonal causal influence patterns provide concrete data-driven behaviors for a virtual peer (or a different learning technology) to emulate for causing an increase/maintenance in curiosity ($+\Delta$) for a group member. The bottom-left and bottom-right corners of Table 4 highlight a comparatively small(er) number of intrapersonal causal influence patterns, those where one's past behavior (within the last minute) leads to an increase/maintenance of curiosity (and vice-versa).

D. Implementation in Pedagogical Agents

Fig. 5. Virtual peer interacts with the customized Outbreak board game.

We have developed a virtual peer that aims to elicit curiosity for young children in a collaborative tabletop game named Outbreak [82]. We applied heuristics derived from our theoretical framework of curiosity as described in this article and in [83], as well as our subsequent computational model of curiosity [81], [84]. The objective was to design social interactions for the virtual peer to perform for evoking key curiosity drives which lead to identification and acquisition of knowledge. First, we developed a virtual peer whose appearance resembles a 9-14 years old child with ambiguous gender and ethnicity (to avoid inadvertent gender and ethnicity stereotypes that might lead to stereotype threat for particular groups of students [85]). Figure 5 illustrates the virtual peer setup during a gameplay session.

The architecture underlying the virtual peer resembles a dialogue system architecture. The first module developed has

TABLE IV
Conditional Granger causality results (at 1% level of significance, 1 minute granularity) between verbal/nonverbal behavior time series and thin-slice curiosity (increase/maintenance) time series. Interpersonal causal influence patterns reflect interdependence of *other's* behavior on one's own curiosity $+\Delta$ (top left), or *other's* curiosity $+\Delta$ on one's own behavior (top right). Intrapersonal causal influence patterns reflect interdependence of *one's own* behavior on curiosity $+\Delta$ (bottom left), or *one's own* curiosity $+\Delta$ on behavior (bottom right).

	Causal Influence (From)	Causal Influence (To)	Standardized Causal Strength (G-ratio)	Causal Influence (From)	Causal Influence (To)	Standardized Causal Strength (G-ratio)
Interpersonal	Confusion-related facial expressions		0.591		Question Asking (on-task)	0.834
	Surprise-related facial expressions		0.420		Head Nod	0.648
	Head Turn		0.303		Justification	0.281
	Head Nod		0.222		Turn Taking Outdegree	0.276
	Justification		0.208		Joy-related facial expressions	0.225
	Task Sentiment (positive)	Curiosity$_{increase/maintain}$	0.198	Curiosity$_{increase/maintain}$	Agreement	0.138
	Idea Verbalization		0.196		Flow-related facial expressions	0.057
	Joy-related facial expressions		0.159		Lateral Head Inclination	0.001
	Evaluation (negative)		0.151			
	Lateral Head Inclination		0.083			
	Question Asking (on-task)		0.046			
	Turn Taking Outdegree		0.025			
Intrapersonal	Uncertainty		1		Confusion-related facial expressions	0.306
	Lateral Head Inclination		0.863		Uncertainty	0.278
	Joy-related facial expressions	Curiosity$_{increase/maintain}$	0.520	Curiosity$_{increase/maintain}$	Task Sentiment (negative)	0.238
	Head Nod		0.134		Head Nod	0.2
	Task Sentiment (negative)		0.021		Task Sentiment (positive)	0.026

been a verbal behavior generation module [86] that allows the virtual peer to produce age-appropriate curiosity-inducing conversational moves including question asking, hypothesis verbalization, argument, and justification during the discussions elicited by the collaborative board game. This generation module allows the virtual peer to generate the kind of sentences that the human-human analyses described above have shown to increase curiosity.

VII. Implications

Few current learning technologies aim to foster the metacognitive factors shown to increase learning (although see [87], [88], [89], [90], [91]). We believe reasons for this include (i) lack of theoretical formalisms and real-time measurement approaches to capture the intricate nature of metacognitive and socio-emotional factors such as creativity, curiosity, grit, help-seeking, self-explanation ability, etc [92], and (ii) dearth of an operational way to embed this theoretical understanding into computational models that can leverage mapping between behaviors and their underlying mechanisms to offer scaffolding [93]. Here, we attempted to go beyond former research on curiosity inference directly from visual and vocal cues (e.g., [61], [94], [60]), by considering underlying mechanisms that link these low-level cues to curiosity, and empirically validating how such cues interact with group collaboration, interpersonal dynamics and conversation moves. Our research approach holds concrete implications for the design of learning technologies, and multimodal learning analytics.

A. Implications for Learning Technologies

For learning technology design, our prior work has demonstrated that one successful approach is to investigate what forms of verbal and nonverbal behaviors and their corresponding functions are good indicators of curiosity - and good elicitors of curiosity - in human-human interaction. This can facilitate the development of adaptive technologies that look for opportunities to use strategies (that is, to be tactical) to scaffold, maintain and evoke curiosity. The underlying infrastructure needed to translate the proposed theoretical framework into teaching tools and technological learning environments should comprise (i) a curiosity perception module for accurate

multimodal behavior sensing using available sensors (e.g., cameras, microphones and other biometric devices) in real-time interaction, (ii) a curiosity reasoner that outlines how a learning technology should act in to support those same functions, and a generation module to translate the intentions generated by the reasoner into observable verbal and nonverbal behaviors for the learning technology to perform.

Prior work on pedagogical agents/robots [41] makes assumptions regarding what observable behaviors can represent lack of curiosity and should therefore result in a move on the part of the learning technology. Here, instead, we relied on close observation of human interaction to extract the moment-to-moment behaviors that occur during periods of curiosity, and amplified the power of those correlations by introducing the theoretically grounded layer of functions that can be fulfilled by behaviors. Below, we expand on how these functions can be scaffolded using strategies and associated tactics in virtual peers, in other technologies, and even simply by teachers or tutors. Specific strategies and tactics are needed because a virtual peer acting curious is not sufficient. Prior literature [95] and our empirical data show that (i) a curious child may not always attempt to increase the curiosity of another child (nor does curiosity in one child always cause curiosity in another) perhaps because it requires increased cognitive and social effort, and (ii) disinterested children may become increasingly cut-off from the core interactions of the group over time, and therefore it's important to find ways to reignite their interest and engagement. In what follows we summarize the strategies and tactics we found in group informal learning contexts. These can be implemented in a virtual peer, and other learning technologies.

1) Strategies for Supporting Functions of Curiosity: We define strategies as moves made by a third party (learning technology or human coach/peer) in the service of facilitating curiosity. This means that strategies serve as vehicles for an influence attempt in the group, directly (e.g., affecting gains/costs) or indirectly (e.g., controlling critical environmental aspects) affecting the behavior of a group member [28]. We believe that strategies should support underlying functions that contribute to curiosity because this enables computer support

to address the root cause of undesirable behavior or provide a reinforcing means for the root cause of desirable behavior. The success of a strategy can be determined concretely by the extent to which it increases thin-slice curiosity in subsequent time-intervals, and specifically by the probability that it leads to the expected behavior(s) by a target child. We propose a categorization of strategies into three clusters in alignment with the functions of curiosity.

Strategies that can facilitate the *knowledge identification* function should help peers realize potential new knowledge to seek. Our empirical data suggested awareness of novel or complex stimuli as a means towards this end. Prior literature also highlights that raising awareness of the conflict between group members can facilitate knowledge identification [96]. Strategies that can facilitate the *knowledge acquisition* function should stimulate critical thinking by helping peers develop new interpretations and consider alternative perspectives, and prevent them from valuing their cohesiveness and relationships with others so much that they avoid conflict and challenging each other's ideas [28]. Literature suggests that provoking group members out of their comfort zone and encouraging them to rethink/defend their responses can be a means towards this end [97]. Strategies that can facilitate the *intensification of knowledge identification and acquisition* function may create a friendly climate, and honor the seeking of knowledge gained through trial and error. Our empirical data in formal learning contexts suggested that when learners were explicitly told to avoid personal criticism and pro-actively grapple with their intuitive ideas, they often opened up to their peers, with a reduced fear of being seen as incompetent, or being excluded. Literature also emphasizes the provision of a supportive and psychologically safe environment to have beneficial effects on accelerated risk-taking [98].

2) Tactics for Exercising a Particular Strategy: We define tactics as particular observable ways in which a strategy can be exercised. The rationale behind the usage of tactics is that certain forms of group interaction (e.g., sharing findings with peers) are more effective for raising curiosity than others. Thus, it is worthwhile to make some arrangements for provoking such interactions, as they may not occur spontaneously. This approach is inspired by the traditional computer-supported collaborative learning literature on scripting [99], however, we propose a more real-time version of such scripting-based approaches. Our interest is rather in adaptive regulation of the small group interaction "on the fly" [100] by continuously comparing the current curiosity level with a target configuration (e.g., if the likelihood of being in high curiosity level in the following time intervals exceeds a certain threshold), and exercising tactics to restore equilibrium whenever there is a discrepancy in the current and target curiosity level. We propose a categorization of tactics falling into each of the three strategy clusters.

Our observations show scarce evidence of children explicitly making moves to exercise the strategy of *knowledge identification*. However, the literature suggests that attentional anchors in the form of contrasting case comparisons can be leveraged by a coach towards this end [101]. By creating paradoxes, such a tactic can help children notice novel features of the stimulus.

Our data do bring to light different tactics children use in the service of exercising the strategy of *knowledge acquisition*. These tactics include challenging a peer's responses (e.g., I don't think this will be really sturdy though), and asking them to make an explicit link between ideas, representations and solution strategies (e.g., what's your evidence for that?). Prior literature also highlights how tactics such as (i) making group members take positions on a big question raised by a task issue and then present reasons and evidence for and against, (ii) encouraging group members who are in conflict to paraphrase each other's position, can be used in service of similar goals [97]. Finally, we observe that tactics children use for exercising the strategy of *intensifying knowledge identification or acquisition* include expressing curiosity-orientation behavior (e.g., looking at stimulus with surprise, expressing interest in individuals and activities using gaze or body orientation), and excitement about solution strategy of a peer. Concrete tactics such as rewarding risk-taking and providing positive feedback for effort may have a positive impact on fulfilling these goals.

B. Implications for Multimodal Learning Analytics

Joining a long tradition of multimodal human behavior analysis in the Learning Sciences (e.g., [74], [102]), and other application areas, Multimodal Learning Analytics (MMLA) [103] offers a useful lens through which to interpret the sensing of, reasoning about and responding to natural human behavior in learning environments, including both verbal and nonverbal devices. In the current work, we developed a theoretical model of curiosity, built arguments and adduced evidence for multimodal indicators and elicitors of curiosity, combined machine learning and human annotation to refine the theoretical model, and ended with a statistical model that tests and validates the theoretical assumptions.

While we have primarily focused on the implications of this work for the implementation of virtual peers, another straightforward implication of our work bears on the tension in MMLA between preserving enough of the complexity of curiosity-driven behavior in the theoretical framework (to obtain valid scientific insight), while at the same time attaining computational feasibility (making sure annotated behavior representations retain enough contextual information and are learnable by machine learning based algorithmic approaches). The part of our analyses related to sensing highlights a potential multimodal fusion solution of capturing and synchronizing different streams of data at the grain size of the ground truth (in this instance thin-slice annotation of curiosity).

Also relevant to MMLA is our iterative theory-and-data-driven approach for capturing not just noisy and intermittent low-level data gathered by physical sensors, but also translating such data into models that describe the underlying psychological states that are correlated with and sometimes cause those observable behaviors. These mid-level phenomena are more generalizable across learning contexts and across metacognitive phenomena. In fact, little work in MMLA has addressed the temporal dimensions of relationships among verbal and nonverbal behaviors, and CTSEM may provide a useful tool in looking at how *series* of behaviors may predict an underlying state such as curiosity.

These MMLA implications bring with them practical challenges that include (i) combating inaccuracy in detecting behavioral episodes because machine learning approaches are not 100% robust, (ii) developing consistent and opportunistic planning algorithms for selecting responding strategies based on current percepts and tying together associated sets of tactics because human behavior is not always rational and dynamic environmental circumstances may affect pre-planned action recipes, (iii) using appropriate evaluation criteria to test the effectiveness of a responding strategy on subsequent behavior in human-technology interaction because people's application of social rules and heuristics from their domain to the domain of machines is often moderated by socially-competent behavior exhibited by the machine [104].

VIII. LIMITATIONS

We must acknowledge some methodological limitations of this work. First, the small sample size limits generalizability. Second, the reliability of thin-slice annotation of ground truth (curiosity) via crowdsourcing platforms can be improved by varying more carefully factors such as the time-scale (granularity) of ratings, rating scale, choice of crowdsourcing workers, and task setup on the crowdsourcing platform. Third, our approach of combining machine annotation with human judgment for annotation of verbal behaviors increases reproducibility, speed and scalability, without compromising on inter-rater reliability. Despite these advantages, going through machine annotated labels and evaluating their accuracy (de-noising) is a different task than if those labels were not there in the first place (meaning that a completely manual annotation approach had been followed). Future work could have some intermediate points during this de-noising process, where the initial inter-rater reliability for human judgment (left column of Table 3) could be re-evaluated for consistency. Fourth, latent variable models used as empirical validation tools in our work are limited by their ability to make causal inferences, especially in cross-sectional datasets [105].

Fifth, we used a crude proxy to infer emotional states from facial landmarks, and future work could adopt complementary predictive modeling approaches [106]. Although facial expressions have the advantage of being observable and recognizable using current computer vision approaches with high accuracy, they can often be polysemous, ambiguous, and be voluntarily camouflaged for social reasons. That is, a smile may mean embarrassment and not happiness, it's unclear where a head nod means agreement or simply that one follows what the other person is saying, and cultural differences include covering facial expressions with one's hand because it is inappropriate to smile in a company (for example). Such subtle distinctions among underlying mechanisms cannot be teased apart. Sixth, it is important to note that despite the existence of learning opportunities, a given student's unwillingness to learn and explore is over-determined. It can stem from multiple sources, such as unawareness of new information that is to be learned, lack of knowledge (competency) and information-seeking skills, lack of environmental support, shyness in talking to peers, and so forth. Therefore, although our research sketched a detailed outline of individual and interpersonal

functions of curiosity, this is only a first step. The next phase of this research will use the virtual peer as a way of assessing the validity of our model by observing the impact of given functions and given behaviors on the human interlocutors.

IX. FUTURE WORK

To examine the effectiveness of the computational model in eliciting curiosity during agent-child interaction, we are developing a semi-automated Wizard-Of-Oz system that integrates (i) the curiosity elicitation reasoner to decide specific real-time contexts during which verbal and non-verbal responses of the virtual peer can elicit curiosity in group work, (ii) the behavior understanding module to understand real-time game and behavior updates of multiple game players. This kind of assessment of the model allows us to turn behaviors on and off to see which ones have the highest impact on curiosity (a luxury not afforded by observing child-child interaction).

We have also begun to extend our analysis of nonverbal behaviors beyond facial expressions to other parts of the body. We are therefore beginning to analyze the nature of hand movements to understand their role in indexing and in triggering curiosity. Hand movements, including manipulative actions and communicative gestures [107], are a key aspect of hands-on collaborative learning. From a distributed cognition perspective, manipulative actions on artifacts and external representations embedded in the physical surroundings help shape people's thoughts [108] by serving important epistemic functions in information gathering and ease of cognitive work [109], and supporting communication [110]. Although hand movements have long been recognized as indispensable to the kinds of physical and epistemic exploration that contribute to curiosity [20], there is a research gap in identifying quantitative relationships between hand movement and curiosity through empirical observation. As an extension of the theoretical framework of curiosity outlined here, we will conduct analyses of key relationships [111] between hand movements and verbal behaviors. As inputs into these analyses, we have articulated a taxonomy of such hand movements with respect to form and meaning. Specifically, we have narrowed down gestures to three kinds of hand movements commonly found in our in-lab study corpus, based on McNeill's well-known gesture coding scheme [112].

X. CONCLUSION

The results, model, and framework presented here are part of a larger research effort to understand the social scaffolding of curiosity and to use that understanding to implement a virtual peer to increase children's curiosity in formal and informal learning contexts, and from there their ability and desire to learn. The theoretical framework presented in this article lays the foundation of a computational model of curiosity that can enable a virtual peer to sense the real-time curiosity level of each member in small group interaction, and the impact of each member's behavior on the others. Here, we articulated key social factors that account for curiosity in learning in social contexts, proposed and empirically validated a novel theoretical framework that disentangles individual and interpersonal functions linked to curiosity and behaviors that fulfill these functions.

The framework connects observable behaviors of children with their underlying curiosity state. Our analyses reveal interpersonal dynamics that indicate as well as elicit curiosity states. As indicators of a curiosity state, the framework reveals strong positive predictive relationships between the interpersonal functions of knowledge identification, acquisition and intensification on curiosity, whereby individual peers had an impact on one another. In terms of eliciting curiosity states, the framework identifies interpersonal behavior patterns that causally influence the increase or maintenance of curiosity. Taken together, these results provide strong evidence for the social nature of curiosity, and the need to disentangle its interpersonal precursors from its individual precursors. In the context of a science game, a pedagogical agent that speaks like a child was also developed to generate the kind of sentences that our current analyses of human-human data have shown to result in increased curiosity. Through designing such learning technologies, we hope to provide additional pedagogical supports to help children develop knowledge-seeking skills driven by intrinsic curiosity.

REFERENCES

[1] D. Park, E. Tsukayama, G. P. Goodwin, S. Patrick, and A. L. Duckworth, "A tripartite taxonomy of character: Evidence for intrapersonal, interpersonal, and intellectual competencies in children," *Contemporary Educational Psychology*, vol. 48, pp. 16–27, 2017.

[2] S. Engel, "Is curiosity vanishing?" *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 48, no. 8, pp. 777–779, 2009.

[3] ——, "Children's need to know: Curiosity in schools," *Harvard Educational Review*, vol. 81, no. 4, pp. 625–645, 2011.

[4] J. Jirout and D. Klahr, "Childrens scientific curiosity: In search of an operational definition of an elusive concept," *Developmental Review*, vol. 32, no. 2, pp. 125–160, 2012.

[5] C. Kidd and B. Y. Hayden, "The psychology and neuroscience of curiosity," *Neuron*, vol. 88, no. 3, pp. 449–460, 2015.

[6] Q. Wu and C. Miao, "Modeling curiosity-related emotions for virtual peer learners," *IEEE Computational Intelligence Magazine*, vol. 8, no. 2, pp. 50–62, 2013.

[7] G. Gordon, C. Breazeal, and S. Engel, "Can children catch curiosity from a social robot?" in *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 2015, pp. 91–98.

[8] E. Law, M. Yin, J. Goh, K. Chen, M. A. Terry, and K. Z. Gajos, "Curiosity killed the cat, but makes crowdwork better," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 2016, pp. 4098–4110.

[9] P.-Y. Oudeyer, "Intelligent adaptive curiosity: a source of self-development," 2004.

[10] J. M. Parr and M. A. Townsend, "Environments, processes, and mechanisms in peer learning," *International journal of educational research*, vol. 37, no. 5, pp. 403–423, 2002.

[11] T. B. Kashdan and F. D. Fincham, "Facilitating curiosity: A social and self-regulatory perspective for scientifically based interventions," *Positive psychology in practice*, pp. 482–503, 2004.

[12] M. E. Jordan and R. R. McDaniel Jr, "Managing uncertainty during collaborative problem solving in elementary school teams: The role of peer influence in robotics engineering activity," *Journal of the Learning Sciences*, vol. 23, no. 4, pp. 490–536, 2014.

[13] Q. Wu, C. Miao, and C. Leung, "Modeling curiosity in virtual companions to improve human learners' learning experience," in *Web Intelligence and Intelligent Agent Technology (WI-IAT), 2015 IEEE/WIC/ACM International Conference on*, vol. 2. IEEE, 2015, pp. 59–66.

[14] J. Cassell, M. Ananny, A. Basu, T. Bickmore, P. Chong, D. Mellis, K. Ryokai, J. Smith, H. Vilhjálmsson, and H. Yan, "Shared reality: physical collaboration with a virtual peer," in *CHI'00 extended abstracts on Human factors in computing systems*. ACM, 2000, pp. 259–260.

[15] W. James, *The principles of psychology*. New York: Henry Holt and Company, 1890.

[16] W. McDougall, *An introduction to social psychology*. Boston : J.W. Luce & co., 1921.

[17] J. Litman, "Curiosity and the pleasures of learning: Wanting and liking new information," *Cognition & emotion*, vol. 19, no. 6, pp. 793–814, 2005.

[18] J. Piaget, *Psychology of intelligence*. New York: Littlefield, Adams, 1969.

[19] J. Hunt, "Motivation inherent in information processing and action," 1963.

[20] D. E. Berlyne, "Conflict, arousal, and curiosity." 1960.

[21] G. Loewenstein, "The psychology of curiosity: A review and reinterpretation." *Psychological bulletin*, vol. 116, no. 1, p. 75, 1994.

[22] R. Golman and G. Loewenstein, "Curiosity, information gaps, and the utility of knowledge," 2015.

[23] O. Spektor-Levy, Y. K. Baruch, and Z. Mevarech, "Science and scientific curiosity in pre-schoolthe teacher's point of view," *International Journal of Science Education*, vol. 35, no. 13, pp. 2226–2253, 2013.

[24] Y. K. Baruch, O. Spektor-Levy, and N. Mashal, "Pre-schoolersverbal and behavioral responses as indicators of attitudes and scientific curiosity," *International Journal of Science and Mathematics Education*, vol. 1, no. 14, pp. 125–148, 2016.

[25] M. i. Zion and I. Sadeh, "Curiosity and open inquiry learning," *Journal of Biological Education*, vol. 41, no. 4, pp. 162–169, 2007.

[26] E. Bonawitz, P. Shafto, H. Gweon, N. D. Goodman, E. Spelke, and L. Schulz, "The double-edged sword of pedagogy: Instruction limits spontaneous exploration and discovery," *Cognition*, vol. 120, no. 3, pp. 322–330, 2011.

[27] B. B. Henderson, "Parents and exploration: The effect of context on individual differences in exploratory behavior," *Child Development*, pp. 1237–1245, 1984.

[28] D. E. Cartwright and A. E. Zander, "Group dynamics research and theory." 1953.

[29] J. Piaget, *The language and thought of the child*. Psychology Press, 1959, vol. 5.

[30] R. Radloff, "Opinion evaluation and affiliation." *The Journal of Abnormal and Social Psychology*, vol. 62, no. 3, p. 578, 1961.

[31] Z. Dörnyei and T. Murphey, *Group dynamics in the language classroom*. Ernst Klett Sprachen, 2003.

[32] P. Van den Bossche, W. H. Gijselaers, M. Segers, and P. A. Kirschner, "Social and cognitive factors driving teamwork in collaborative learning environments: Team learning beliefs and behaviors," *Small group research*, vol. 37, no. 5, pp. 490–521, 2006.

[33] Y. Engeström, R. Engeström, and M. Kärkkäinen, "Polycontextuality and boundary crossing in expert cognition: Learning and problem solving in complex work activities," *Learning and instruction*, vol. 5, no. 4, pp. 319–336, 1995.

[34] U. Cress and J. Kimmerle, "A systemic and cognitive view on collaborative knowledge building with wikis," *International Journal of Computer-Supported Collaborative Learning*, vol. 3, no. 2, p. 105, 2008.

[35] G. Hatano and K. Inagaki, "Intellectual curiosity," *Cyuko Shinsho (in Japanese)*, 1973.

[36] H. Ogata and Y. Yano, "Combining knowledge awareness and information filtering in an open-ended collaborative learning environment," *International Journal of Artificial Intelligence in Education (IJAIED)*, vol. 11, pp. 33–46, 2000.

[37] S. Reiss, "Multifaceted nature of intrinsic motivation: The theory of 16 basic desires." *Review of general psychology*, vol. 8, no. 3, p. 179, 2004.

[38] J. M. Keller, "Strategies for stimulating the motivation to learn," *Performance Improvement*, vol. 26, no. 8, pp. 1–7, 1987.

[39] S. B. Shum and R. D. Crick, "Learning dispositions and transferable competencies: pedagogy, modelling and learning analytics," in *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*. ACM, 2012, pp. 92–101.

[40] A. C. Graesser, N. K. Person, and J. P. Magliano, "Collaborative dialogue patterns in naturalistic one-to-one tutoring," *Applied cognitive psychology*, vol. 9, no. 6, pp. 495–522, 1995.

[41] Q. Wu, X. Han, H. Yu, Z. Shen, and C. Miao, "The innovative application of learning companions in virtual singapura," in *Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems*. International Foundation for Autonomous Agents and Multiagent Systems, 2013, pp. 1171–1172.

[42] Q. Wu and C. Miao, "Curiosity: From psychology to computation," *ACM Computing Surveys (CSUR)*, vol. 46, no. 2, p. 18, 2013.

[43] W. M. Wundt, *Grundzüge de physiologischen Psychologie*. W. Engelman, 1874, vol. 1.

[44] Standard aligned system. [Online]. Available: http://www.pdesas.org/

[45] D. OConnor, "Application sharing in k-12 education: Teaching and learning with rube goldberg," *TechTrends*, vol. 47, no. 5, pp. 6–13, 2003.

[46] J. Ritchie, J. Lewis, C. M. Nicholls, R. Ormston *et al.*, *Qualitative research practice: A guide for social science students and researchers*. Sage, 2013.

[47] M. T. Chi and R. Wylie, "The icap framework: Linking cognitive engagement to active learning outcomes," *Educational Psychologist*, vol. 49, no. 4, pp. 219–243, 2014.

[48] D. L. Schwartz and T. Martin, "Inventing to prepare for future learning: The hidden efficiency of encouraging original student production in statistics instruction," *Cognition and Instruction*, vol. 22, no. 2, pp. 129–184, 2004.

[49] T. Sinha, M. Kapur, R. West, M. Catasta, M. Hauswirth, and D. Trninic, "Differential benefits of explicit failure-driven and success-driven scaffolding in problem-solving prior to instruction." *Journal of Educational Psychology*, 2020.

[50] D. W. Johnson and R. T. Johnson, "Energizing learning: The instructional power of conflict," *Educational Researcher*, vol. 38, no. 1, pp. 37–51, 2009.

[51] B. Renner, "Curiosity about people: The development of a social curiosity measure in adults," *Journal of personality assessment*, vol. 87, no. 3, pp. 305–316, 2006.

[52] D. Heylen, "Investigating mind markers in design meetings," in *Group Decision and Negotiation (GDN) 2006: International Conference, Karlsruhe, Germany, June 25-28, 2006; Proceedings*. KIT Scientific Publishing, 2006, p. 132.

[53] M. R. Luce and S. Hsi, "Science-relevant curiosity expression and interest in science: an exploratory study," *Science Education*, vol. 99, no. 1, pp. 70–97, 2015.

[54] B. McDaniel, S. D'Mello, B. King, P. Chipman, K. Tapp, and A. Graesser, "Facial features for affective state detection in learning environments," in *Proceedings of the Cognitive Science Society*, vol. 29, no. 29, 2007.

[55] N. Ambady and R. Rosenthal, "Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis." 1992.

[56] D. J. Follmer, R. A. Sperling, and H. K. Suen, "The role of mturk in education research: Advantages, issues, and future directions," *Educational Researcher*, vol. 46, no. 6, pp. 329–334, 2017.

[57] T. Sinha and J. Cassell, "We click, we align, we learn: Impact of influence and convergence processes on student learning and rapport building," in *Proceedings of the 1st Workshop on Modeling INTER-PERsonal SynchRONy And infLuence*. ACM, 2015, pp. 13–20.

[58] R. Zhao, T. Sinha, A. W. Black, and J. Cassell, "Socially-aware virtual agents: Automatically assessing dyadic rapport from temporal patterns of behavior," in *International Conference on Intelligent Virtual Agents*. Springer, 2016, pp. 218–233.

[59] J. Kruger, U. Endriss, R. Fernández, and C. Qing, "Axiomatic analysis of aggregation methods for collective annotation," in *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*. International Foundation for Autonomous Agents and Multiagent Systems, 2014, pp. 1185–1192.

[60] B. Nojavanasghari, T. Baltrušaitis, C. E. Hughes, and L.-P. Morency, "Emoreact: a multimodal approach and dataset for recognizing emotional responses in children," in *Proceedings of the 18th ACM International Conference on Multimodal Interaction*. ACM, 2016, pp. 137–144.

[61] S. D. Craig, S. D'Mello, A. Witherspoon, and A. Graesser, "Emote aloud during learning with autotutor: Applying the facial action coding system to cognitive–affective states during learning," *Cognition and Emotion*, vol. 22, no. 5, pp. 777–788, 2008.

[62] J. Cassell and T. Bickmore, "Negotiated collusion: Modeling social language and its relationship effects in intelligent agents," *User modeling and user-adapted interaction*, vol. 13, no. 1, pp. 89–132, 2003.

[63] R. Zhao, T. Sinha, A. W. Black, and J. Cassell, "Automatic recognition of conversational strategies in the service of a socially-aware dialog system," in *17th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, 2016, p. 381.

[64] Q. V. Le and T. Mikolov, "Distributed representations of sentences and documents." in *ICML*, vol. 14, 2014, pp. 1188–1196.

[65] R. Farkas, V. Vincze, G. Móra, J. Csirik, and G. Szarvas, "The conll-2010 shared task: learning to detect hedges and their scope in natural language text," in *Proceedings of the Fourteenth Conference on Computational Natural Language Learning—Shared Task*. Association for Computational Linguistics, 2010, pp. 1–12.

[66] R. Swanson, B. Ecker, and M. Walker, "Argument mining: Extracting arguments from online dialogue," in *Proceedings of the 16th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, 2015, pp. 217–226.

[67] P. Jansen, N. Balasubramanian, M. Surdeanu, and P. Clark, "Whats in an explanation? characterizing knowledge and inference requirements for elementary science exams."

[68] S. Negi, "Suggestion mining from opinionated text," *ACL 2016*, p. 119, 2016.

[69] L. Dong, F. Wei, Y. Duan, X. Liu, M. Zhou, and K. Xu, "The automated acquisition of suggestions from tweets." in *AAAI*, 2013.

[70] J. Andreas, S. Rosenthal, and K. McKeown, "Annotating agreement and disagreement in threaded discussion." in *LREC*. Citeseer, 2012, pp. 818–822.

[71] B. Cahour, "Characteristics, emergence and circulation in interactional learning," *Affective learning together*, pp. 52–70, 2013.

[72] J. F. Grafsgaard, K. E. Boyer, R. Phillips, and J. C. Lester, "Modeling confusion: facial expression, task, and discourse in task-oriented tutorial dialogue," in *International Conference on Artificial Intelligence in Education*. Springer, 2011, pp. 98–105.

[73] T. Baltrušaitis, P. Robinson, and L.-P. Morency, "Openface: an open source facial behavior analysis toolkit," in *Applications of Computer Vision (WACV), 2016 IEEE Winter Conference on*. IEEE, 2016, pp. 1–10.

[74] K. Ryokai, C. Vaucelle, and J. Cassell, "Literacy learning by storytelling with a virtual peer," in *Proceedings of the conference on computer support for collaborative learning: foundations for a CSCL community*. International Society of the Learning Sciences, 2002, pp. 352–360.

[75] D. Gatica-Perez, L. McCowan, D. Zhang, and S. Bengio, "Detecting group interest-level in meetings," in *Acoustics, Speech, and Signal Processing, 2005. Proceedings.(ICASSP'05). IEEE International Conference on*, vol. 1. IEEE, 2005, pp. I–489.

[76] B. Schuller, R. Müller, F. Eyben, J. Gast, B. Hörnler, M. Wöllmer, G. Rigoll, A. Höthker, and H. Konosu, "Being bored? recognising natural interest by extensive audiovisual integration for real-life application," *Image and Vision Computing*, vol. 27, no. 12, pp. 1760–1774, 2009.

[77] C. Lai, J. Carletta, S. Renals, K. Evanini, and K. Zechner, "Detecting summarization hot spots in meetings using group level involvement and turn-taking features." in *INTERSPEECH*, 2013, pp. 2723–2727.

[78] C. C. Driver, J. H. Oud, and M. C. Voelkle, "Continuous time structural equation modelling with r package ctsem," *Journal of Statistical Software*.

[79] J. Förster and F. Strack, "Influence of overt head movements on memory for valenced words: a case of conceptual-motor compatibility." *Journal of personality and social psychology*, vol. 71, no. 3, p. 421, 1996.

[80] M. Ding, Y. Chen, and S. L. Bressler, "Granger causality: basic theory and application to neuroscience," *Handbook of time series analysis: recent theoretical developments and applications*, pp. 437–460, 2006.

[81] T. Sinha, Z. Bai, and J. Cassell, "Curious minds wonder alike: Studying multimodal behavioral dynamics to design social scaffolding of curiosity," in *Data Driven Approaches in Digital Education: 12th European Conference on Technology Enhanced Learning, EC-TEL 2017, Tallinn, Estonia, September 12–15, 2017, Proceedings*, É. Lavoué, H. Drachsler, K. Verbert, J. Broisin, and M. Pérez-Sanagustín, Eds. Cham: Springer International Publishing, 2017, pp. 270–285.

[82] A. To, J. Holmes, E. Fath, E. Zhang, G. Kaufman, and J. Hammer, "Modeling and designing for key elements of curiosity: Risking failure, valuing questions," *Transactions of the Digital Games Research Association*, vol. 4, no. 2, 2018.

[83] T. Sinha, Z. Bai, and J. Cassell, "A new theoretical framework for curiosity for learning in social contexts," in *Data Driven Approaches in Digital Education: 12th European Conference on Technology Enhanced Learning, EC-TEL 2017, Tallinn, Estonia, September 12–15, 2017,*

Proceedings, É. Lavoué, H. Drachsler, K. Verbert, J. Broisin, and M. Pérez-Sanagustín, Eds.   Cham: Springer International Publishing, 2017, pp. 254–269.

[84] B. Paranjape, Z. Bai, and J. Cassell, "Predicting the temporal and social dynamics of curiosity in small group learning," in *International Conference on Artificial Intelligence in Education*.   Springer, 2018, pp. 420–435.

[85] J. R. Shapiro and A. M. Williams, "The role of stereotype threats in undermining girls and womens performance and interest in stem fields," *Sex Roles*, vol. 66, no. 3-4, pp. 175–183, 2012.

[86] B. Paranjape, Y. Ge, Z. Bai, J. Hammer, and J. Cassell, "Towards automatic generation of peer-targeted science talk in curiosity-evoking virtual agent," in *International Conference on Intelligent Virtual Agents*. ACM, 2018, pp. 71–78.

[87] C. Conati and K. Vanlehn, "Toward computer-based support of meta-cognitive skills: A computational framework to coach self-explanation," *International Journal of Artificial Intelligence in Education (IJAIED)*, vol. 11, pp. 389–415, 2000.

[88] V. Aleven, B. Mclaren, I. Roll, and K. Koedinger, "Toward meta-cognitive tutoring: A model of help seeking with a cognitive tutor," *International Journal of Artificial Intelligence in Education*, vol. 16, no. 2, pp. 101–128, 2006.

[89] R. Azevedo, A. M. Witherspoon, A. Chauncey, C. Burkett, and A. Fike, "Metatutor: A metacognitive tool for enhancing self-regulated learning." in *AAAI Fall Symposium: Cognitive and Metacognitive Educational Systems*, 2009.

[90] M. Pifarre and R. Cobos, "Promoting metacognitive skills through peer scaffolding in a cscl environment," *International Journal of Computer-Supported Collaborative Learning*, vol. 5, no. 2, pp. 237–253, 2010.

[91] M. Bannert and P. Reimann, "Supporting self-regulated hypermedia learning through prompts," *Instructional Science*, vol. 40, no. 1, pp. 193–211, 2012.

[92] A. L. Duckworth and D. S. Yeager, "Measurement matters: Assessing personal qualities other than cognitive ability for educational purposes," *Educational Researcher*, vol. 44, no. 4, pp. 237–251, 2015.

[93] R. Reiter-Palmon, T. Sinha, J. Gevers, J.-M. Odobez, and G. Volpe, "Theories and models of teams and groups," *Small Group Research*, vol. 48, no. 5, pp. 544–567, 2017.

[94] A. Baranes, P.-Y. Oudeyer, and J. Gottlieb, "Eye movements reveal epistemic curiosity in human observers," *Vision research*, vol. 117, pp. 81–90, 2015.

[95] D. M. Morrison and V. Rus, "–moves, tactics, strategies, and metastrategies: Defining the nature of human pedagogical interaction," *Design Recommendations for Intelligent Tutoring Systems*, p. 217.

[96] D. R. Forsyth, *Group dynamics*.   Cengage Learning, 2009.

[97] R. Correnti, M. K. Stein, M. S. Smith, J. Scherrer, M. McKeown, J. Greeno, and K. Ashley, "Improving teaching at scale: Design for the scientific measurement and learning of discourse practice," *Socializing Intelligence Through Academic Talk and Dialogue. AERA*, 2015.

[98] A. Edmondson, "Psychological safety and learning behavior in work teams," *Administrative science quarterly*, vol. 44, no. 2, pp. 350–383, 1999.

[99] F. Fischer, I. Kollar, H. Mandl, and J. M. Haake, *Scripting computer-supported collaborative learning: Cognitive, computational and educational perspectives*.   Springer Science & Business Media, 2007, vol. 6.

[100] P. Jermann and P. Dillenbourg, "Group mirrors to support interaction regulation in collaborative problem solving," *Computers & Education*, vol. 51, no. 1, pp. 279–296, 2008.

[101] L. Alfieri, T. J. Nokes-Malach, and C. D. Schunn, "Learning through case comparisons: A meta-analytic review," *Educational Psychologist*, vol. 48, no. 2, pp. 87–113, 2013.

[102] T. Koschmann and C. LeBaron, "Learner articulation as interactional achievement: Studying the conversation of gesture," *Cognition and instruction*, vol. 20, no. 2, pp. 249–282, 2002.

[103] P. Blikstein, "Multimodal learning analytics," in *Proceedings of the third international conference on learning analytics and knowledge*. ACM, 2013, pp. 102–106.

[104] J. Cassell and A. Tartaro, "Intersubjectivity in human–agent interaction," *Interaction studies*, vol. 8, no. 3, pp. 391–410, 2007.

[105] B. Nagengast and U. Trautwein, "The prospects and limitations of latent variable models in educational psychology," *Handbook of educational psychology*, pp. 48–58, 2015.

[106] E. Sariyanidi, H. Gunes, and A. Cavallaro, "Automatic analysis of facial affect: A survey of registration, representation, and recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 37, no. 6, pp. 1113–1133, 2015.

[107] V. I. Pavlovic, R. Sharma, and T. S. Huang, "Visual interpretation of hand gestures for human-computer interaction: A review," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no. 7, pp. 677–695, 1997.

[108] E. Hutchins, "The distributed cognition perspective on human interaction," *Roots of human sociality: Culture, cognition and interaction*, vol. 1, p. 375, 2006.

[109] D. Kirsh and P. Maglio, "On distinguishing epistemic from pragmatic action," *Cognitive science*, vol. 18, no. 4, pp. 513–549, 1994.

[110] J. Cassell, D. McNeill, and K.-E. McCullough, "Speech-gesture mismatches: Evidence for one underlying representation of linguistic and nonlinguistic information," *Pragmatics & cognition*, vol. 7, no. 1, pp. 1–34, 1999.

[111] P. Fournier-Viger and V. S. Tseng, "Mining top-k non-redundant association rules," in *International Symposium on Methodologies for Intelligent Systems*.   Springer, 2012, pp. 31–40.

[112] D. McNeill, *Hand and mind: What gestures reveal about thought*. University of Chicago press, 1992.

Tanmay Sinha is a postdoctoral researcher in Learning Sciences and Technology Design at ETH Zurich. His doctoral work focused on studying the pedagogical value of deliberate, guided failure to foster students' conceptual understanding and transfer. He is interested in the role of socio-emotional and interpersonal factors in learning, and more generally, in how scaffolded human and technology interventions may open new opportunities for students to fully reap the benefits of failure-driven exploration.

Zhen Bai is an Assistant Professor in the Department of Computer Science at University of Rochester. She is interested in designing embodied and intelligent interfaces that support lifelong learning and quality of life for people with diverse abilities and backgrounds by augmenting social interaction and meaning making in a playful and collaborative manner.

Justine Cassell is SCS Dean's Professor in the School of Computer Science at Carnegie Mellon University, and until recently was Director of the Human-Computer Interaction Institute in the School of Computer Science. She is currently on leave from Carnegie Mellon University to hold the founding international chair at PRAIRIE Paris Institute on Interdisciplinary Research in AI (one of French President Macron's four new AI Institutes), and hold the associated position of Directrice de Recherche at Inria Paris. Her research focuses on understanding natural forms of communication, and then creating technological tools for those forms of communication and linguistic expression to flourish in the digital world.